



XX 26-SEP-2000; 2000EP-0308421.  
 XX  
 PF  
 PR 29-SEP-1999; 99US-0156602.  
 XX  
 XX (PFIZ.) PFIZER PROD INC.  
 PA  
 XX  
 XX King KW, Madura RA, Rosey EL;  
 PI  
 DR WPI: 2001-309781/33.  
 XX N-PSDB; AAS03285.  
 XX  
 XX New apoprotein antigens encoded by mhp3 gene from Mycoplasma  
 PT hyopneumoniae useful as a vaccine for treating or preventing diseases  
 PT caused by Mycoplasma hyopneumoniae -  
 PT  
 PS  
 PS Claim 11; Page 17-18; 38pp; English.

| Query Match | Similarity | 99.0%  | Score     | 2134       | DB | 225    | Length | 451  |
|-------------|------------|--|-----------|------------|----|--------|--------|------|
| Best Local  | Similarity | 99.8%  | Pred. No. | 8.6e-163   |    |        |        |      |
| Matches     | 419        | Conservative   | 0         | Mismatches | 1  | Indels | 0      | Gaps |
| OY          | 2          | WDEFTTKEEKSDADNONKQITDVSKTISGLVNERKSEITMAAKADANKHFGNMAIVTAGTV    | 61        |            |    |        |        |      |
| Db          | 30         | Wdkettkeeksadnqkqitdvsksisglvnerkseimaakadankhfglnmatvtagtv      | 89        |            |    |        |        |      |
| OY          | 62         | NONSNSOGMEAIIOOLGATGTGETISVDSSTAELGKSYSLANTNKNVWVLGSGFOHDAP      | 121       |            |    |        |        |      |
| Db          | 90         | nonsnsgmeaiaioolgatgtgetisvdsstaeltgksyslantnknvwlsgfghgdat      | 149       |            |    |        |        |      |
| OY          | 122        | TRWLKIPENKQLETERKNIILIGIDWDTFNNVPTGCRYIMLYKTEBAGWLAGYANASFLA     | 181       |            |    |        |        |      |
| Db          | 150        | trwlkipenkqftekhnlllgidwtdtcenvlptgrylnllykteeagylagyanasfla     | 209       |            |    |        |        |      |
| OY          | 162        | KRFPPDPYTKRSIIVGGGSPAVTDFIAGYLAGIKRAMNLKNSDKKTKITTDKRIEMAGD      | 241       |            |    |        |        |      |
| Db          | 210        | Krfppdptkrsaiivgggspravtdfiagylagikawmlknsdkktktittdkrielnlgid   | 269       |            |    |        |        |      |
| OY          | 242        | VODTTEKLELEIAKDKRSTLLAVAGPLTELEFSDIIANQNRVLYIGVPTDQSLVYTKRK      | 301       |            |    |        |        |      |
| Db          | 270        | vodttekelleleiaekdkrptcllavagprtlefstdlaqnndrlyllygvtldqsllyttkk | 329       |            |    |        |        |      |
| OY          | 302        | NKEFTSILKNLGYSVFSVLSDLVYTKSNSRNLAGFEFGKRSATVYLGIDRFVADIATSL      | 361       |            |    |        |        |      |
| Db          | 330        | nkeftsilknlgysvfsvlsdlytkssnsrnlagfeygkrsatvylgldrfvdiadta       | 389       |            |    |        |        |      |
| OY          | 362        | EGNDKKLNTAELASEKKEFEETKTIIPAEVAKTLEIPMPKOPKOPKOESLDKLTIDIN       | 421       |            |    |        |        |      |
| Db          | 390        | egndkkllataelaseekfeektktipaeavrtleipmpakpdpkqgesldkltidn        | 449       |            |    |        |        |      |

|    |   |
|----|---|
| XX |   |
| DT | 07-SEP-2001 (first entry)   |
| XX |   |
| DE | Mycoplasma hyopneumoniae recombinant MHP3 antigen.  |
| XX |   |
| KW | MHP3; antigen; vaccine; enzootic mycoplasma pneumonia; mutant;<br>antibody; immunoassay; immunotherapy; anti-idiotypic antibody; mutein.  |
| XX |   |
| OS | Mycoplasma hyopneumoniae.   |
| FH | Key Location/Qualifiers   |
| FT | Misc-difference 70  |
| FT | /note= "Encoded by RGA"   |
| FT | Misc-difference 417..422  |
| FT | /note= "Encoded by AATTACCGATAT"  |
| XX |   |
| PN | EPI090995-A2.   |
| XX |   |
| PD | 11-APR-2001.  |
| XX |   |
| PF | 26-SEP-2000; 2000EP-0308421.  |
| XX |   |
| PR | 29-SEP-1999; 99US-0156602.  |
| XX |   |
| PA | (PFIZ ) PFIZER PROD INC.  |
| XX |   |
| PI | King KW, Madura RA, Rosey EJ;   |
| XX |   |
| DR | WPJ: 2001-309781/33.  |
| DR | N-PSDB; AAS03286.   |
| XX |   |
| PT | New aporettein antigens encoded by mhp3 gene from Mycoplasma<br>hyopneumoniae useful as a vaccine for treating or preventing diseases<br>caused by Mycoplasma hyopneumoniae -   |
| XX |   |
| PS | Claim 4; Page 19-21; 38pp; English.   |
| XX |   |
| CC | The sequence represents Mycoplasma hyopneumoniae a recombinant MHP3<br>antigen lacking the first 28 amino acids (the putative signal sequence).<br>MHP3 antigen and its fragments are useful in manufacturing a vaccine for<br>treating or preventing a disease or disorder in an animal, especially<br>pig, caused by M. hyopneumoniae infection e.g. enzootic mycoplasma<br>pneumonia. The mhp3-encoded proteins may be used as immunogens to<br>generate antibodies which immunospecifically bind such an immunogen.<br>The antibodies generated against the antigen are useful in diagnostic<br>immunoassays, passive immunotherapy and generation of anti-idiotypic<br>antibodies. Mhp3 proteins may also be used in immunoassays, e.g. to<br>detect or measure in a biological sample from a vaccinated or<br>potentially infected test animal the presence of antibodies to the<br>antigen, and thus to monitor the immune response and/or to diagnose<br>infection of the animal. |
| XX |   |
| SO | Sequence 423 AA;  |

|                       |        |   |         |              |
|-----------------------|--------|---|---------|--------------|
| Query Match           | 98.4%  | Score 2120  | DB 2120 | Length 423   |
| Best Local Similarity | 100.0% | Pred. No. 1e-161  |         |              |
| Matches               | 416    | Conservative  | 0       | Mismatches 0 |
|                       |        |   | Indels  | 0            |
|                       |        |   | Gaps    | 0            |
| QY                    | 1      | MMDKETETKEKADNONKOITDVSKISGLVNERKSEIIMAAKADANKHFLGMAIYTAGCT   | 60      |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
| Db                    | 1      | mmddcttceeksaadqmqitdvskisglvnerkseimaakadankhiflmaivtagt     | 60      |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
| QY                    | 61     | VNDNSFNOSGWEAIDOLGALTGETTSDVSS7AELEGKRYSSLANNKVMWVLSGFQHGDA   | 120     |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
| Db                    | 61     | vndnsfnsgweaialqilgaligteltvsdstaelegkrysslanenkmvwvlsqfghda  | 120     |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
| QY                    | 121    | FRWMLKPEENKOLFPEKNIIILGIDMTPEENVIPIRGRIINLTKTEPFGAGTAGVNAASLT | 180     |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
| Db                    | 121    | ftrwfkipeenkqlfexnlilgldwtdeemvipgrgynlntkceegwalsgynaaatl    | 180     |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
| QY                    | 181    | AKKPPSDPKKSAIYIIGGISPAVDTPLAGYLAGIKAMNKNKSDKKTKITTDKIEINLGE   | 240     |              |
|                       |        |   |         |              |
|                       |        |   |         |              |
|                       |        |   |         |              |

Db 181 akfipspdktrsaivigggispavtdfiagylagikawmlknsdkklttdkieinlgt 240  
 QY 241 DVODSTKERLEQIASKDKPSTLLAVAGPTEIFSDIIANQNDRYLIGVTDQSLVYTKT 300  
 Db 241 dvodstkerleqiaaskdkpstllavaagptelfsdianqndrylignvtdqslvytk 300  
 QY 301 KKEFTSILKNLIGSVFVSLDLYTKKSNRLAGEFEGKSAVYVYLGIKDRFVIADTS 360  
 Db 301 knkfesilknlgysvfasvlsdlytkksnarnlagfegkksatvylgikdrfviadts 360  
 QY 361 LENDKKLATEAISEAKKEFEKTKTTPAEVAKTLEIPEMPKQPKQKQESIDKL 416  
 Db 361 legndkklateaieakkefektktipaeevtktleipempkqpkqgesidkl 416  
 RESULT 3  
 AAR40856  
 ID AAR40856 standard; Protein; 453 AA.  
 AC AAR40856;  
 XX  
 DT 07-MAR-1994 (first entry)  
 XX  
 KM 43kd regression associated antigen.  
 KM Regression associated antigen; tumour; immunotherapy;  
 XX anti-idiotypic antibodies; antibodies; tumour regression.  
 OS Mycoplasma hyorhinis.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 80 /note= "Tryptophan encoded by TGA, normal in  
 FT Mycoplasma hyorhinis."  
 FT Misc-difference 124 /note= "Tryptophan encoded by TGA, normal in  
 FT Mycoplasma hyorhinis."  
 FT Misc-difference 165 /note= "Tryptophan encoded by TGA, normal in  
 FT Mycoplasma hyorhinis."  
 FT Misc-difference 344 /note= "Tryptophan encoded by TGA, normal in  
 FT Mycoplasma hyorhinis."  
 XX  
 PN US5242823-A.  
 PD 07-SEP-1993.  
 XX  
 PF 07-MAR-1986; 86US-0837494.  
 XX  
 PR 07-MAR-1986; 86US-0837494.  
 PR 16-SEP-1987; 87US-0097910.  
 PR 11-DEC-1987; 87US-0131815.  
 PR 04-JAN-1988; 88US-0138923.  
 PR 16-MAR-1990; 90US-0474730.  
 PR 02-OCT-1992; 92US-0956546.  
 XX  
 PA (ITGE-) INT GENETIC ENG INC.  
 XX  
 PI Fareed GC, Ghosh-dastidar P, Jar-how L, Sen A;  
 XX  
 DR WPI; 1993-295229/37.  
 DR N-PSDB; AAQ47816.  
 XX  
 PT DNA encoding a regression-associated antigen from M. hyorhinis -  
 PT is used to obtain prods. for diagnosis, localisation and therapy  
 PT of tumours  
 XX  
 PS Disclosure; Figure 3; 40pp; English.  
 XX  
 CC Regression associated antigens (RA's) are identified in material  
 CC from neoplastic cells by their immunological reactivity with  
 CC regression associated antibodies from the serum of patients

CC diagnosed as undergoing regression of a tumour. RA's can be used  
 CC for tumour immunotherapy and for producing and purifying antibodies  
 CC which can be used for tumour diagnosis, localisation and therapy.  
 CC The antibodies can also be used for the production of  
 CC anti-idiotypic antibodies which can also be used in immunotherapy.  
 XX  
 SQ Sequence 453 AA;  
 Query Match 32.6%; Score 702; DB 14; Length 453;  
 Best Local Similarity 38.8%; Pred. NO. 6.7e-48;  
 Matches 170; Conservative 82; Mismatches 146; Indels 40; Gaps 14;  
 QY 5 ETTKEKSAONONKQITDVSKISGLVNERKSEIIMAAKADANKHGLMMAIVTAGTVNDN 64  
 Db 28 etdkegkll-----rld-----nsfvdrqaeleka-----knfdftvlltagtvqdk 73  
 QY 65 SPNOSGWEAT---QOLGALTG-GEITSVDSSTAELRGKYSILANTKNVWVLSGFOHD 119  
 Db 74 sfngsiweavlehydqlektlnldrvsqetnngseligkyknfngknkvwlitgfgqg 133  
 QY 120 APTRMUKIPEN----KQLETEKNIILGIDW--TPTEENVIPGRVINTYKTEEAGWL 171  
 Db 134 efpkflqtdsngkkyssdllaekkvliavawdwliskcdkdlkagnfisllykteegtl 193  
 QY 172 AGYANASFLAKKPPSDPTKRSATVIGGCIAPAVDTFAGYLAGIKAMNLNSDKTKITTY 231  
 Db 194 ayyasakflaykfpndeakrtlpfsggngagvdfdiagflaglakynndnpakvlsd 253  
 QY 232 DKIEINLGFVODSTKERLEQIASKDKPSTLLAVAGPTEIFSDII--ANQNDRYLIGV 289  
 Db 254 nlnldtqf--lsndktatfnglyvns--slvjpyvsgltsvvdalksnkdkyllgv 310  
 QY 290 DTQDSLYTKTKNKEFSSILKNLIGSVFVSLDLYTKKSNRLAGEFEGK---KSAIV 345  
 Db 311 dtqdskltpa-tvftfslehlgrtyevltdvllkedsrkfsgfsltpanav 369  
 QY 346 YLGIKDRFVADIATSLBGNDKILATEAISEAKKEFEKTKTTPAEVAKTLEIPEM--PD 403  
 Db 370 ykgsldfvgvsnstvdaadvkagelinaetadfkxqanpc-nyksvlgpmtlnd 428  
 QY 404 KQDPKQESIDKLTIDIN 421  
 Db 429 ndaknekalneiklin 446  
 RESULT 4  
 AAR67582  
 ID AAR67582 standard; Protein; 432 AA.  
 XX  
 AC AAR67582;  
 XX  
 DT 06-OCT-1995 (first entry)  
 XX  
 DE Cancer metastasis protein.  
 XX  
 KM Cancer metastasis; clinical tests; detection.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 55 /note= "corresponding codon TGA"  
 FT Misc-difference 99 /note= "corresponding codon TGA"  
 FT Misc-difference 140 /note= "corresponding codon TGA"  
 FT Misc-difference 319 /note= "corresponding codon TGA"  
 FT Misc-difference 319 /note= "corresponding codon TGA"  
 FT Misc-difference 420 /note= "corresponding codon TGA"  
 FT Misc-difference 424 /note= "corresponding codon TGA"  
 FT Misc-difference 424 /note= "corresponding codon TGA"

```

XX JP06319559-A.
XX
XX 22-NOV-1994.
XX
XX 12-MAY-1993; 93JP-0144165.
XX
XX 12-MAY-1993; 93JP-0144165.
XX
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX WPI: 1995-040317/06.
XX
XX N-PSDB; AAO79124.
XX
XX DNA encoding polypeptide involved in cancer metastasis - useful
XX PT for the study of the mechanism of metastasis and in clinical
XX PT tests
XX
XX Claim 1; Page 2; 13pp; Japanese.
XX
XX AAO79124 encodes AAR67582 a protein involved in cancer metastasis,
XX CC which may be used for the study of metastatic mechanisms, and for
XX CC clinical tests to determine the presence or absence of cancer
XX CC metastasis.
XX
XX SQ Sequence 432 AA:

Query Match 32.1%; Score 692.5; DB 16; Length 432;
Best Local Similarity 39.0%; Pred. No. 3.6e-47;
Matches 168; Conservative 81; Mismatches 141; Indels 41; Gaps 14;

OY 5 ETTKEKSDNKNKQITDVSKISGLVNERKSEIMAKADANKHFGILMAIYTAGTVNDN 64
DB 3 eckdegkll-----flfd-----nsfvkdrqaeleka-----knfdntvlltaagtvqdk 48
OY 65 SFNOSGWEAI-----QOLGALTG-GEITVSSTALEGKYSILANTNKNWVWLSGFQHD 119
DB 49 sfngsiweavevlehnydqlektlnldrvsgeltnqseligkyknflngknkvlltfgf99g 108
OY 120 AFTRWLKIPEIN-----KOLFEKNIILIGIDW---TDTEENVIPTRGYINLTKTEEAGWL 171
DB 109 efpkflkqtdsngkkysdliaekkvliavdwldskedkdlkaghflsilykteeagfl 168
OY 172 AGYANASFLAKKFPSPDPTKRSAIVIGGISPAPVDFIAGYLAGIRAMWLKNSDKRTKITT 231
DB 169 agyaasskfilykfnpdeakrtlapfg9ghgagvtldflaglakynndnptgkvltisd 228
OY 222 DKIEINLGFVDVDTSTKERLEOIAISKDKRSTLLAVAGPLTEIFSDII--ANONDRYLIGV 289
DB 229 mnlndtgrf-lsndktatfinglvnks--slvlpvagsltsavdalikkshdtkyllgv 285
OY 290 DTDOSLVYTKTKNKPFETSIILKNLGYSVFSVLSDLVTKKSNRNLAGFERGK---KSATV 345
DB 286 dtdgsklifspa-tvfftslekhgrlyevltdlwlkkesdflgsfsfkltnaanatv 344
OY 346 YLGIKDRFVDIADTSLLEGNDRKLAIEAISEAKKEFEKTKITPAEVEVKTLEIPEM---P 402
DB 345 ykgladddfgvsnstvadadkvkageflneatadfkkgqgnpt-nyksvlgipmlind 403
OY 403 DKOPPKQOESL 413
DB 404 ndakdnkaasl 414

RESULT 5
AAP93343
ID AAP93343 standard; protein; 448 AA.
XX
XX AAP93343;
AC
XX
XX 01-AUG-1990 (first entry)
XX

```

```

DE Gene encoding the 43 kd regression-associated antigen (RA Ag) of
DE Mycoplasma hyorhinis.
XX
XX Tumour regression-associated antigens (RA Ag); in vivo imaging;
XX KW therapy monitoring; cancer therapy; Mycoplasma hyorhinis.
XX
XX Mycoplasma hyorhinis.
XX
XX OS
XX
XX Key Location/Qualifiers
XX FT Peptide 1..25
XX FT Protein 26..448
XX
XX EP308265-A.
XX
XX 22-MAR-1989.
XX
XX 16-SEP-1988; 88EP-0308625.
XX
XX 11-DEC-1987; 87US-0131815, US-097910.
XX
XX (ITGE-) INT GENETIC ENG INC.
XX
XX Fared GC, Sen A, Ghosh-Dastidar P, Liu A, Lee JH.
XX
XX WPI: 1989-087638/12.
XX
XX N-PSDB; AAN90684.
XX
XX Tumour regression-associated antigens and antibodies -
XX PT used in diagnostic tests, monitoring course of therapy and for
XX PT therapy in cancer patients
XX
XX PS Disclosure; : 56pp; English.
XX
XX Regression-associated antigens may be used in diagnostic tests, eg in
XX CC vivo imaging, for monitoring the course of therapy or for therapeutic
XX CC purposes, eg active immunisation protocols in cancer patients or drug
XX CC delivery systems by binding the drug to monoclonal or monospecific
XX CC polyclonal Ab showing specific immunoreactivity with the Ag.
XX
XX SQ Sequence 448 AA:

Query Match 29.8%; Score 642.5; DB 10; Length 448;
Best Local Similarity 38.0%; Pred. No. 3.9e-43;
Matches 166; Conservative 83; Mismatches 145; Indels 43; Gaps 18;

OY 5 ETTKEKSDNKNKQITDVSKISGLVNERKSEIMAKADANKHFGILMAIYTAGTVNDN 64
DB 28 eckdegkll-----flfd-----nsfvkdrqaeleka-----knfdntvlltaagtvqdk 73
OY 65 SFNOSGWEAI-----QOLGALTG-GEITVSSTALEGKYSILANTNKNWVWLSGFQHD 119
DB 74 sfngsiweavevlehnydqlektlnldrvsgeltnqseligkyknflngknkvlltfgf99g 133
OY 120 AFTRWLKIPEIN-----KOLFEKNIILIGIDW---TDTEENVIPTRGYINLTKTEEAGWL 171
DB 134 efpkflkqtdsngkkysdliaekkvliavdwldskedkdlkaghflsilykteeagfl 193
OY 172 AGYANASFLAKKFPSPDPTKRSAIVIGGISPAPVDFIAGYLAGIRAMWLKNSDKRTKITT 231
DB 194 agyaasskfilykfnpdeakrtlapfg9ghgagvtldflaglakynndnptgkvltisd 253
OY 222 DKIEINLGFVDVDTSTKERLEOIAISKDKRSTLLAVAGPLTEIFSDII--ANONDRYLIGV 289
DB 254 mnlndtgrf-lsndktatfinglvnks--slvlpvagsltsavdalikkshdtkyllgv 310
OY 290 DTDOSLVYTKTKNKPFETSIILKNLGYSVFSVLSDLVTKKSNRNLAGFERGK---KSATV 345
DB 311 dtdgsklifspa-tvfftslekhgrlyevltdlwlkkesdflgsfsfkltnaanatv 367
OY 346 YLGIKDRFVDIADTSLLEGNDRKLAIEAISEAKKEFEKTKITPAEVEVKTLEIPEM-PDK 404
DB 368 ykglisd-flgvsnstvadadkvkagefln-etadfkkgqgnpt-nyksvlgipmlindn 424

```

OY 405 OPDKQGESLDTITDIN 421  
Db 425 dakhmekalneldkkin 441

## RESULT 6

AAV05332  
ID AAV05332 standard; peptide; 428 AA.

AC AAV05332;

DT 25-JUN-1999 (first entry)

DE Inflammatory cytokine inducer.

KW Inflammatory cytokine inducer; thrombocytopenia; therapy.

OS Synthetic.

PN JP11089582-A.

XX 06-APR-1999.

PF 19-SEP-1997; 97JP-0273649.

PR 19-SEP-1997; 97JP-0273649.

XX (CHUS) CHUGAI PHARM CO LTD.

PA WPI: 1999-281057/24.

DR N-PSDB; AAX3847.

XX New inflammatory cytokine inducer gene and polypeptide - useful for treatment of thrombocytopenia

PS Claim 3; Page 16-17; 22pp; Japanese.

CC This sequence represents the inflammatory cytokine of the invention. The inflammatory cytokine can be used in a drug, which is useful for the treatment of thrombocytopenia.

XX Sequence 428 AA;

Query Match 17.6%; Score 379.5; DB 20; Length 428;

Best Local Similarity 26.7%; Pred. No. 4, 6e-22; Mismatches 156; Indels 71; Gaps 17;

Matches 115; Conservative 88; Mismatches 156; Indels 71; Gaps 17;

31 NERKSEIMAAKADANKH-----FGLNMAIVTAGTVNDNSFNQSGEAI 74

27 ndesnlstfkekdskyttnangkvvnaelkikpylltdegklddsfnsgafel 86

OY 75 QOLGALTGEITISVDSSTLEBEGKISSLANTKNWNVLSGFQHDFTKMLKIPENKOLF 134

Db 87 kelnkqtlgelnvps-snfesaynsalsagkhivlnghqgsikgy--ldahneel 143

OY 135 TEKNITILGIDMTDFENVIPTRGYNLUTKTEBAGWLAGANSPILAKRPPDPTKRSAT 194

Db 144 etnqkllidldf-dleeykw--fyslqfinkesafitcgyalaswseq--deskriva 197

OY 195 VTGGISPAVTFEIAAGYLAGIKAMNLSNDKKTI-TTDKIEINLGFVDSTKTERLEQ 253

Db 198 sfsggsfpvvtfnegfakglllynnqkh--kskkyhtspvkidsqf-----tagekmt 250

OY 254 IAS-----KDRPSYLLAVAGPLTEIFSDIINONDRYLIGVTDOSLVTYTKTKN 303

Db 251 vlnnvlsstpadkynphvllsvaspat--felvrlankqyivgodsqgm--qdkdr 306

OY 304 FPTSILKNIGSYSVFVSLSDL-----YTKKSNSRNLAGFEFGKKSATVYLGIRDRF 353

Db 307 lltsvlkhkhqavetlidlillekeegykyrvvkkdkkadkshfqtg-----kek 358

OY 354 VDIADTSLGENDKKL-ATEAISSEAKKEFEKTKTIPAEVYRKTLEIPMPD-KOPDKOE 411  
Db 359 igvaenhfsnteegaklnkikaikmfke----lpedfvykynsdkaikdgkldnvse 414

OY 412 SLDKLITDIN 421

Db 415 rleallsaln 424

## RESULT 7

AAW22727  
ID AAW22727 standard; Protein; 429 AA.

AC AAW22727;

DT 26-FEB-1998 (first entry)

DE Membrane protein M161Ag.

KW M161Ag; membrane protein; leukaemia; treatment.

OS Homo sapiens.

PN JP09157295-A.

PD 17-JUN-1997.

PF 05-DEC-1995; 95JP-0344504.

PR 05-DEC-1995; 95JP-0344504.

PA (KAGA-) KAGAKU GIYUTSU SHINKO JIGYODAN.

DR WPI: 1997-369470/34.

XX N-PSDB; AAT75133.

XX Membrane protein M161Ag - useful in development of leukaemia treating agent

PS Claim 1; Fig 1; 6pp; Japanese.

CC This sequence is a membrane protein designated M161Ag. Oligonucleotide probes for the cDNA were designed from putative N-terminal M161Ag peptide sequences isolated from P39 cells. M161Ag can be used in the development of a leukaemia treating agent.

XX Sequence 429 AA;

Query Match 17.3%; Score 372; DB 18; Length 429;

Best Local Similarity 26.8%; Pred. No. 1, 9e-21; Mismatches 167; Indels 64; Gaps 18;

Matches 117; Conservative 88; Mismatches 167; Indels 64; Gaps 18;

13 ADNOKOIT---DVSKISGLVNERKSEIMAAKADANKHFGLNMAIVTAGTVNDNSFNQ 68

Db 26 andesnlstfkekdskyttnangkvvnael-----lklypvlltdegklddsfnq 80

OY 69 SGWEAIGQLGALTGEITISVDSSTLEBEGKISSLANTKNWNVLSGFQHDFTKMLKIP 128

Db 81 satealalnkhqtlgelnvps-snfesaynsalsagkhixvlnghqgsikgy--ld 137

OY 129 ENKOLFTEKNIIILGIDWTDTEENVIPGKRYINLFTKTEBAGWLAVANASFLAKKPPSDP 188  
 Db 138 ahreelernqlk1lgldf-dleleykx--fyslqfnikesaltcyalasxseq---de 191  
 OY 189 TKRSALVIGGSPAVTDFIAGYIKAMNKNKSDKRTKI-TTDKIEINLGFEDVDNST 247  
 Db 192 srtvvaafvggafpvtctfnegfakypnqh--ksaklyhtspvrydsqf-----ta 244  
 OY 248 KERLEQIAS-----KDRPSTLLAVAGPLTEIFSDIINQNDRLIGVTDQSLVY 297  
 Db 245 gekmvtlnnvlastpadykynphvllvavpact--fevtlpmkgvyigydsdgml- 301  
 OY 298 TKTAKKFTSILKNGYSFVSLSL-----YTKKSNRLNAGFEFGKKSATVYL 347  
 Db 302 -gdkdrilltsvllkhkgvayellql1lekeegykyrvvkdkaakksahfgtq----- 354  
 OY 348 GKDRFVFIADTSLGNDKRL-ATEAISPAKKEFEKRTTIPAEVRKLELPEMPD-KQ 405  
 OY 355 --kexxiyvaenqfnteeqaklnmkleaklmfke----lpedivkylnsdkaikdgnk 408  
 OY 406 PDKQOESLDKLITTDIN 421  
 Db 409 ldnveerleallsaaln 424

## RESULT 8

AAV81632  
 ID AAV81632 standard; Protein: 351 AA.

AC AAV81632;

DT 24-MAY-2000 (first entry)

XX Streptococcus pneumoniae type 4 protein sequence #132.

KW Streptococcus pneumoniae; vaccine; screening; protein antigen;  
 KW antibacterial; antinflammatory; meningitis; infection; diagnosis;  
 KW pneumococcal disease.

XX Streptococcus pneumoniae.

XX WO200006737-A2.

PN 10-FEB-2000.

PD 27-JUL-1999; 99WO-GB02451.

XX 27-JUL-1998; 98GB-0016337.

PR 19-MAR-1999; 99US-0125164.

XX (MICR-) MICROBIAL TECHNICS LTD.

XX Gilbert CFG, Hansbro PM;

XX WPI: 2000-195300/17.

XX New Streptococcal protein, useful as a vaccine, for diagnosis of  
 PT pneumococcal diseases and for screening agents capable of antagonizing  
 PT or inhibiting expression of the protein

XX Claim 1; Page 85; 108pp; English.

XX AAV81501 to AAV81679 represent specifically claimed protein sequences  
 CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent  
 CC specifically claimed nucleotide sequences isolated from S. pneumoniae.  
 CC The sequences have antibacterial and antinflammatory properties.  
 CC The protein sequences, and fragments of them, are useful as immunogens  
 CC and/or antigens. The nucleotide sequences can be used in vaccines and in  
 CC diagnostic assays. The proteins and nucleotides can be useful for the  
 CC detection and diagnosis of S. pneumoniae. The protein sequences are also  
 CC useful for screening an agent capable of antagonizing, inhibiting or  
 CC interfering with the function or expression of the proteins in which the  
 CC agent is useful for treatment or prophylaxis of S. pneumoniae infection

CC and meningitis. AAA05591 to AAA05614 represent primers used in the  
 CC exemplification of the present invention.  
 XX  
 XX Sequence 351 AA;

Query Match 9.6%; Score 206; DB 21; Length 351;  
 Best Local Similarity 25.5%; Pred. No. 2.9e-08;  
 Matches 98; Conservative 57; Mismatches 146; Indels 84; Gaps 19;

OY 33 RKSEIMAKKADANKHFGNMAIVTAGTVNDSEFNQSGMEALQQLGA---LTGGEITVSD 89  
 Db 25 rrsrnaasssdvkv---tkaalvtctgyvddksfnqsawegiqavgkehnlslskdngficyf 80  
 OY 90 SSTAALE-----GKYSLSLANTNNKWWVLSCFQHDADFTRLMKIPENKOLFTEKNII 140  
 Db 81 qltseadyannlqgaagynllfgy-----gfalmaa-----vdaakehdldnyf 126  
 OY 141 ILGIDWTDTEENVIPGKRYINLFTKTEBAGWLAVANASFLAKKPPSDPTKRSALVIGGI 200  
 Db 127 llddvlykdkqv-----asvlfadnesgylagvaak-----tlckekvgfvggl 171  
 OY 201 -SPAVTDFIAGYIKAMNKNKSDKRTKITDKEINLGFVYDTSKERLEQIASDK 259  
 Db 172 esevlsrteagfkagv-----asvdpslkvydvag-sfg-daakgkclaaagyaagad- 223  
 OY 260 PSTLLAVAGPL-TEIFSDIINQNDR-----YLIGVTDQSL--VYT---KTKNKFPT 306  
 Db 224 -lyvqvaggygagvfaeakslnesrpenekwvlygvdrdgaagkykskdgkeanfylv 281  
 OY 307 SLKKNLGYSVFSVLSDLTYTKKSNRNLAGFEFGKKSATVYLGIKDRFYDIADTSLGNDK 366  
 Db 282 sltkvggtlvxdl-----snkaergefp9gq--Vlyvlykdkgydlatvtnlseegk 330  
 OY 367 KLTEAISEAKKEFEKRTTIPAE 391  
 Db 331 k-----avedakaklidsvkvpeks 351

## RESULT 9

AAW55066  
 ID AAW55066 standard; Protein: 328 AA.

AC AAW55066;

DT 02-OCT-1998 (first entry)

XX Streptococcus pneumoniae SP007 protein.

XX Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;  
 KW detection; pneumonia; otitis media; meningitis.

XX Streptococcus pneumoniae.

XX WO9818930-A2.

XX 07-MAY-1998.

XX 30-OCT-1997; 97WO-US19422.

XX 31-OCT-1996; 96US-0029960.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Chol GH, Hromocky A, Johnson LS, Kunsch CA;

XX WPI: 1998-272224/24.

XX N-PSDB; AAV27326.

XX Nucleic acid encoding antigenic peptide(s) from Streptococcus  
 PT pneumoniae - or their epitope-containing fragments, useful in  
 PT protective or therapeutic vaccines, and for diagnosis







Db 180 vvidrfqagfeygv-adaakejgkeitvdt-----kyaasfadpa 218  
Qy 262 TLLAVAGPLTEIFSDII-----ANONDR-YLIGVDTQSL--VYF 298  
Db 219 kgkalaamnygvdfifnasgatggvfgaeakdlneqsgdkvwvlygdrdqdagky- 277  
Qy 299 KTK-----NKFTSLIKNLGYSVFSLDLYTKKSRLNLAGFEFGKKSATVYLKIDRF 333  
Db 278 kkkdsgkedfllstlkygtavgdianaleak-----fpygehlvy-gikdgs 326  
Qy 354 VDIADSLSGNDKILATEAISEAKKEFEKTKTIPAEVRYKLTLEIPDMQ 403  
Db 327 vldtqyl--ndkt-----keavktakdvksagdkvpdkpe 361  
RESULT 13  
AAU07352  
ID AAU07352 standard; Protein; 166 AA.  
XX  
XX AAU07352;  
XX  
XX 03-JAN-2002 (first entry)  
XX  
XX Interleukin-X (IL-X).  
XX  
XX Interleukin-X; IL-X; autostimulatory factor; Epstein Barr virus; EBV;  
XX  
XX lymphoblastoid; B cell.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX FT Misc-difference 150  
XX FT Misc-difference 160 /note= "Encoded by TGA"  
XX FT Misc-difference 160 /note= "Encoded by TGA"  
XX FT  
XX PM US2001019713-A1.  
XX PD 06-SEP-2001.  
XX PF 16-JAN-2001; 2001US-0760541.  
XX PR 08-AUG-1994; 94US-0287387.  
XX PR 11-FEB-1993; 93US-0026132.  
XX  
XX (BLAZ/) BLAZAR B A.  
XX PA (WEBB/) WEBB A C.  
XX PI Blazar BA, Webb AC;  
XX  
XX WPI; 2001-570181/64.  
XX N-PSDB; AAS11798.  
XX  
XX New IL-X polypeptides useful e.g. for enhancing the growth of cells in  
XX vitro, for growing primary human B cell cultures to facilitate  
XX proliferation of these cell lines, as therapeutics or prophylactics, as  
XX molecular weight standards -  
XX  
XX Claim 1; Page 7; 16pp; English.  
XX  
XX The invention relates to a novel secreted autostimulatory factor,  
XX interleukin-X (IL-X), from an Epstein Barr virus (EBV)-carrying  
XX lymphoblastoid cell line. IL-X proteins can be used to enhance the growth  
XX of cells in vitro, or to grow primary human B cell cultures to facilitate  
XX proliferation of these cell lines, as therapeutics or prophylactics, as  
XX molecular weight standards, as inert proteins in an assay, or in the  
XX detection of antibodies that are immunoreactive with IL-X. Nucleotide  
XX sequences encoding the peptides could be used to produce recombinant  
XX peptides or they could be used as probes or primers for diagnostic  
XX and/or analytical polymerase chain reaction (PCR) procedures, or as DNA  
XX or RNA sizing standards. The present sequence represents the amino acid  
XX sequence of IL-X.

SQ Sequence 166 AA;  
Query Match 8.0%; Score 173; DB 22; Length 166;  
Best Local Similarity 27.3%; Pred. No. 4, 2e-06;  
Matches 50; Conservative 37; Mismatches 56; Indels 40; Gaps 8;  
Qy 197 GGGISPAVDFIAGIAGIKAMNLKNSDKKTKI-TTDKIEINLGFVQDSTRERLEQIA 255  
Db 2 9ggafeyvttfnegfakgallynqkh--ksklyhspvldsfg-----tagkmtvl 54  
Qy 256 S-----KDPSTLLAVAGPLTEIFSDIIANONDRYLIGVDTQSLVYTKTKNKF 305  
Db 55 nvlstpadvkvphvllsvagpat--fetvrlankgyvlgvdadqgm--gdkdrll 110  
Qy 306 TSLIKNLGYSVFSLDLSL-----YTKNSNRNLGAFEGKKSATVYLKIDRF 355  
Db 111 tsvikhkhqavetllidlllekeegykpvyvkkdkkdkkwshtlgt-----kekwi 162  
Qy 356 IAD 358  
Db 163 vae 165  
RESULT 14  
AAR33280  
ID AAR33280 standard; Protein; 341 AA.  
XX  
XX AAR33280;  
XX  
XX 17-DEC-2001 (updated)  
XX DT 11-MAY-1993 (first entry)  
XX  
XX P39-beta.  
XX  
XX Polymerase chain reaction; PCR; primer: open reading frame; ORF; P39;  
XX antigen; 39-alpha; 39-beta; serum; human; Lyme patient; p39-beta;  
XX p39-alpha; promoter; -10; -35; ribosomal binding site; RBS.  
XX  
XX Borrelia burgdorferi.  
XX OS  
XX  
XX USN7664731-N.  
XX PN  
XX 01-DEC-1992.  
XX PD  
XX 05-MAR-1990; 90US-0487716.  
XX PF  
XX 05-MAR-1990; 90US-0487716.  
XX PR  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
XX PA  
XX Schwan TG, Simpson WJ;  
XX PI  
XX WPI; 1993-017799/02.  
XX DR N-PSDB; AAO34672.  
XX  
XX New Borrelia burgdorferi protein and DNA isolates - used in  
XX detection of Lyme borreliosis infection and in prodn. of  
XX antibodies and vaccines  
XX  
XX Disclosure; Page 36-39; 69pp; English.  
XX  
XX The sequences given in AAR31013 and AAR33280 represent the B.  
XX burgdorferi proteins P39-alpha and P39-beta. The DNA encoding these  
XX proteins was isolated by polymerase chain reaction (PCR) using the  
XX primer sequences given in AAO34673-76. This DNA contained two open  
XX reading frames (ORF). It was therefore concluded that the P39 antigen  
XX which has been previously described is not one protein but two, 39-alpha  
XX and 39-beta. The P39 signal appears to be enhanced when both genes  
XX are present. Gene 1 encodes a 339 amino acid protein with a  
XX calculated molecular weight of 36,926 kD. The protein encoded by this  
XX gene which is reactive with serum from human Lyme patients. The ORF  
XX of gene 2 has been designated p39-beta. This genes' ORF starts 116



Search completed: July 18, 2002, 15:32:20  
Job time: 187 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:29:55 ; Search time 33.42 Seconds  
(without alignments)  
1253.726 Million cell updates/sec

Title: US-09-676-249A-4  
Perfect score: 2135  
Sequence: 1 MMDETTKEKSKADNOMKQI.....KQPKQKQSLDKLTIDNNL 423

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID     | Description          |
|------------|-------|-------------|--------|----|--------|----------------------|
| 1          | 552   | 25.6        | 461    | 2  | B90555 | ABC transporter xy   |
| 2          | 203.5 | 9.4         | 350    | 2  | F86804 | basic membrane pro   |
| 3          | 202   | 9.4         | 350    | 2  | G95097 | lipoprotein [impor   |
| 4          | 198   | 9.2         | 374    | 2  | C97865 | conserved hypotet    |
| 5          | 192   | 8.9         | 357    | 2  | D96986 | probable lipoprote   |
| 6          | 181   | 8.4         | 353    | 1  | H71340 | membrane lipoprote   |
| 7          | 180.5 | 8.4         | 350    | 1  | C70009 | ABC transporter (l   |
| 8          | 178.5 | 8.3         | 359    | 2  | F72418 | basic membrane pro   |
| 9          | 177   | 8.2         | 341    | 2  | E70147 | basic membrane lip   |
| 10         | 173   | 8.0         | 357    | 2  | AH1610 | CD4+ T cell-stimul   |
| 11         | 168.5 | 7.8         | 316    | 2  | C82946 | hypothetical prote   |
| 12         | 168   | 7.8         | 357    | 2  | AD1248 | CD4+ T cell-stimul   |
| 13         | 165.5 | 7.7         | 360    | 2  | H70147 | basic membrane pro   |
| 14         | 162.5 | 7.5         | 525    | 2  | C82914 | conserved hypotet    |
| 15         | 161   | 7.5         | 339    | 2  | F70147 | basic membrane pro   |
| 16         | 152.5 | 7.1         | 591    | 2  | D64204 | membrane lipoprote   |
| 17         | 150   | 7.0         | 524    | 2  | D82944 | hypothetical membr   |
| 18         | 146   | 6.8         | 353    | 2  | G70147 | basic membrane pro   |
| 19         | 140   | 6.5         | 379    | 2  | H75318 | membrane lipoprote   |
| 20         | 139.5 | 6.5         | 337    | 2  | H79734 | Delonococcus radiod  |
| 21         | 139.5 | 6.5         | 337    | 2  | AH2591 | membrane lipoprote   |
| 22         | 136   | 6.3         | 547    | 2  | E29504 | mercuryl (IT) reduct |
| 23         | 134   | 6.2         | 539    | 2  | D82886 | conserved hypotet    |
| 24         | 132   | 6.1         | 326    | 2  | G95857 | hypothetical prote   |
| 25         | 131.5 | 6.1         | 1558   | 2  | B71603 | RESA-H3 antigen pr   |
| 26         | 128   | 5.9         | 657    | 2  | S73428 | probable lipoprote   |
| 27         | 126   | 5.8         | 349    | 2  | F84246 | hypothetical prote   |
| 28         | 125.5 | 5.8         | 763    | 2  | A82863 | hypothetical prote   |
| 29         | 123.5 | 5.8         | 1223   | 2  | E88451 | protein K10D2.1 [1   |

|    |       |     |      |   |        |                     |
|----|-------|-----|------|---|--------|---------------------|
| 30 | 123.5 | 5.7 | 626  | 2 | C25035 | colicin Ia - Esche  |
| 31 | 122.5 | 5.7 | 322  | 2 | F84236 | ABC transporter (l  |
| 32 | 122.5 | 5.7 | 2346 | 2 | T13829 | Tpr homolog - fruit |
| 33 | 120   | 5.6 | 1546 | 2 | G90603 | lipoprotein [impor  |
| 34 | 119.5 | 5.5 | 384  | 2 | I40867 | hypothetical prote  |
| 35 | 119.5 | 5.5 | 2285 | 2 | T12796 | probable transglyc  |
| 36 | 118.5 | 5.5 | 556  | 2 | H82301 | peptide ABC transp  |
| 37 | 118.5 | 5.5 | 1140 | 2 | S73786 | hypothetical prote  |
| 38 | 118.5 | 5.5 | 1963 | 2 | B98002 | IGA-specific metal  |
| 39 | 118   | 5.5 | 1031 | 2 | C81302 | probable type I sl  |
| 40 | 117   | 5.4 | 798  | 2 | S62791 | probable lipoprote  |
| 41 | 117   | 5.4 | 1959 | 2 | AG1085 | hypothetical prote  |
| 42 | 116.5 | 5.4 | 396  | 2 | H97500 | hypothetical prote  |
| 43 | 116.5 | 5.4 | 396  | 2 | AD2719 | conserved hypotet   |
| 44 | 115.5 | 5.4 | 350  | 2 | F70139 | exported protein (  |
| 45 | 115   | 5.3 | 282  | 2 | C97271 | probable xylanase/  |

## ALIGNMENTS

## RESULT 1

B90555  
ABC transporter xylase-binding lipoprotein [imported] - Mycoplasma pulmonis (strain U  
C/Species: Mycoplasma pulmonis  
C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C/Accession: B90555  
R:Chamand, I.; Hellig, R.; Ferris, S.; Barde, V.; Samson, D.; Galisson, F.; Moszer,  
Nucleic Acids Res. 29, 2145-2153, 2001  
A/Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p  
A/Reference number: A99512; MID:21267165; PMID:11353084  
A/Accession: B90555  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1461 <KUR>  
A/Cross-references: GB:AL445566; PID:g14089760; PIDN:CAC13519.1; GSPDB:GN00153  
A/Experimental source: strain UAB CTIP  
C/Genetics:  
A/Gene: MYPU\_3460  
A/Genetic code: SGC3

| Query Match   | 25.6% | Score 552         | DB 2 | Length 461 |
|---|-------|-------------------|------|------------|
| Best Local Similarity   | 31.2% | Pred. No. 5.1e-27 |      |            |
| Matches 133; Conservative 82; Mismatches 161; Indels 50; Gaps 11;       |       |                   |      |            |
| OY 13 ADONKQIT--DVSKISGLVNER-----KSEIMAAKADANKHGLMMAITAGCTV 61          |       |                   |      |            |
| DB 28 AONPNKTNNSMLDSKIKIDLSQKEVETETOKIVENKIKQASLETOK----VLTITADGNI 82   |       |                   |      |            |
| OY 62 NDNFSNGWEAIOQLALNGEITS-----VDSFALEEGKYSLSANTNK 107                |       |                   |      |            |
| DB 83 DKSFSNOQVYESQKTLDPFDKATKSONKEAENOKHDKNDYNSAVKLEQNYKALDRGY 142     |       |                   |      |            |
| OY 108 NWVYLSEFOHGDAFTWRLKIPENKQKTEKNIILIGIDWTDEN-VIPGRYINLYKTE 166     |       |                   |      |            |
| DB 143 TTWILTFQOQNEIENFLNDENNLRPEKKNVKIIGVDMAPNANSKIPOGSLISLEKTE 202    |       |                   |      |            |
| OY 167 EAGLAIYANASFLAKFPSPDPTKRSALVIGGIGSPAVTDFIAGYLAGICIAWN--LKNSD 224 |       |                   |      |            |
| DB 203 EAGQAOQVYASDAGTGYANNAEKRAISAFGGGDFAGVTFDLNFFGIGIANWSEAMEN 262    |       |                   |      |            |
| OY 225 KRTITTDKIEINIGFVODPTSTKERIEQJASKKPSLTLAVAGPLREIFSDIT---ANQ 281   |       |                   |      |            |
| DB 263 KKVATIVENLVLDGF--IPNAEKNEVSNVYETKSTISLPVACPFPGVAVDVYAKRTSD 321   |       |                   |      |            |
| OY 282 NDRYLIGVDTDQSLVYTKRKNKFFTSILKNLGYSFVSLDLYT-----KKSNSRNL 333      |       |                   |      |            |
| DB 322 EDRIYIVGVDTDQSLFNDNSKRFFTSIVKNIAFPVYQIILALLTKRDESVYLKRGNDKFL 381 |       |                   |      |            |
| OY 334 AGFEPKKSATVYIGIDRFVADIADTSLGENDKRLAIEAISEA--KKEFEKTTITPAEE 391   |       |                   |      |            |
| DB 382 GS--NPKNLVLKRGLSAKFVNITKSRVKESIKTQADTSTOKALDKWANNPNNSKRIKEM 438  |       |                   |      |            |

OY 392 VKTLE 397  
DB 439 TNGDLE 444

RESULT 2  
F86804  
basic membrane protein A [Imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: F86804  
R:Boletín, A.; Winkler, P.; Mauger, S.; Jallón, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A:Reference number: AB6625; MUID:21235186; PMID:11337471  
A:Accession: F86804  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-350 <STO>  
A:Cross-references: GB:AE005176; PID:g12724428; PIDN:AAK05536.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: bmpA

Query Match 9.4%; Score 203.5; DB 2; Length 350;  
Best Local Similarity 24.7%; Pred. No. 1.9e-05;  
Matches 97; Conservative 63; Mismatches 157; Indels 75; Gaps 17;

OY 20 INDVKGISLVNERKSEIMAAKADANKHGLMAIVTAGTVNDNSFNQSGWEALQOLGA 79  
DB 13 LASVAVLAGCRSHDAAGSGAKTD-----LKAATVTEIGVNDSSFNQSAWEGLOQWCK 66  
OY 80 LTG-----GEITSVDSSTAELGKYSLLANTKNVVLGSGFHGDATFRLKIPENKOLF 134  
DB 67 ENNLKKGCTYTFQNSNASDYTNNSAQCGYKLLFGFSLQDATSAAR---NN--- 120  
OY 135 TEKNTIIGIDMTDENVIPTGRYINLTFTKEAGWLAGYANASFLAKKFPSPDKRSAT 194  
DB 121 PKNFVIYDSVYKQKNV-----ASATFADNESAYLAGVAAR-----ATKTKNI 165  
OY 195 -VIGGIGSPAVTDFLAGYLAGIKAMNLSKDKTKITTDKIINIGFVODPSTKERLEQ 253  
DB 166 GFTGGMQSDVITREFEKGTEAGAKSVN-----PDIKVDVQYAGSFSDAAKKTIAA 215  
OY 234 IASKDKPSTLLAVAGPL-TEIFSDIIA-----NQNDR-YLIGVDTQSLV--YT---KT 300  
DB 216 AMYGAGDDVYVGCAGGVGVSEAKALNSTKNEADKWVIGVDDQDELTCKYKSKDKD 275  
OY 301 KKKFTSLIKNLGYSVSVSLDYTKKNSRNLAG--FEFGKKSATVYLGIKDRFVIDAD 358  
DB 276 SNFVLVSTIKVGV---NVVKDIAKTKDKGKPGGTIVYDILKNGVNLGL-----DSAN 326  
OY 359 TSLGNDKKLATEATSEAKKEFEKTKTTPAE 390  
DB 327 SEIK-----DAVAKAKADIIDGKITVPSK 350

RESULT 3  
G95097  
lipoprotein (Imported) - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
C:Accession: G95097  
R:Teitelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: G95097

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-374 <KUR>  
A:Cross-references: GB:AE005672; PIDN:AAK74976.1; PID:g14972319; GSPDB:GN00164; TIGR:  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP0845

Query Match 9.4%; Score 202; DB 2; Length 350;  
Best Local Similarity 25.7%; Pred. No. 2.4e-05;  
Matches 98; Conservative 55; Mismatches 145; Indels 84; Gaps 19;

OY 33 RKSEIMAAKADANKHFGIMMAIVTAGTVNDNSFNQSGWEALQOLGA---LTGGEITSVD 89  
DB 25 RSSRNAASSDVK-----TKAAIVDTGVDKSFQNSAMEGLQAWGKEHNSKDNGLTFYF 80  
OY 90 SSTABLE-----GKYSLLANTKNVWVLGSGFHGDATFRLKIPENKOLFTEKNIT 140  
DB 81 OSTSEADYANMLQQAAGSYNLTFGV-----GFLNNA-----VKDAKKEHTDLNLYV 126  
OY 141 ILGIMDTDENVIPTGRYINLTFTKEAGWLAGYANASFLAKKFPSPDKRSATYIGGGI 200  
DB 127 LIDDIYKQKNV-----ASVTFADNESGYLAGVAAR-----TTKQVGFVGGI 171  
OY 201 -SPAVTDFIAGYLAGIKAMNLSKDKTKITTDKIEMIGFDVODPSTKERLEQIASKDK 259  
DB 172 ESEVISREFEAGKAGV-----ASVDPSTIKQVDYAG-SFG-DAAKGKTIAAQVAAGAD- 223  
OY 260 PSTLLAVAGPL-TEIFSDIINQNR-----YLIGVTDQSL-VYF---KTKNKKFT 306  
DB 224 --TVQVAGGTGAGVFAAKSLNESRPENKVVYGVDDQEAEGKYTSKDKGSNFVLV 281  
OY 307 SILKNLAGSVFSVSLDYTKKNSRNLAGFEFGKKSATVYLGIKDRFVIDADTSLEGNDK 366  
DB 292 STLKQVGTIVDI-----SNKAERGEFPGQ---YIVSLDKGVADIATNLSEBCK 330  
OY 367 KIATFAISEAKKEFEKTKTTP 388  
DB 331 K---AVEDAKAKIIDGSVKVP 348

RESULT 4  
C97965  
conserved hypothetical protein spr0747 [Imported] - Streptococcus pneumoniae (strain  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001  
C:Accession: C97965  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; Dehoff, B.S.  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Balaz, R.H.; Jaskunas, S.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: C97965  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-374 <KUR>  
A:Cross-references: GB:AE007317; PIDN:AAK99551.1; PID:g15458340; GSPDB:GN00174  
C:Genetics:  
A:Gene: spr0747

Query Match 9.2%; Score 198; DB 2; Length 374;  
Best Local Similarity 25.4%; Pred. No. 4.6e-05;  
Matches 97; Conservative 54; Mismatches 147; Indels 84; Gaps 18;

OY 33 RKSEIMAAKADANKHFGIMMAIVTAGTVNDNSFNQSGWEALQOLGA---LTGGEITSVD 89  
DB 49 RSSRNAASSDVK-----TKAAIVDTGVDKSFQNSAMEGLQAWGKEHNSKDNGLTFYF 104  
OY 90 SSTABLE-----GKYSLLANTKNVWVLGSGFHGDATFRLKIPENKOLFTEKNIT 140

```

Db      105  OSTSADYAAANNLQQAAGSYNLI-----FQVGFALHNAVEBAKHEHTDLNVY 150
Qy      141  ILGIDMTDEVVITPGRINILTYKTEAGWLAGYANASFLAKKFPSPDTKSAIYIGGI 200
Db      151  LIDDVKKOKKVV-----ASYTFADNESGYLAGVAAAAC-----TTKTKOVGFVGII 195
Qy      201  -SPAATDIEIAGYGIKAMWLKNSDKKRTITTDKEINLGFDPYOOTSTKRELEQJASDK 259
Db      196  ESEYISREAPKFGV-----ASVDPSTIKVOYDIAG-SFG-DAAKGTITAAQYIAGAD- 247
Qy      260  PSTLLAVAGPL-TEIFSDIILANONDR-----YLIGVDTQSL--VYT-----KTKNKEFT 306
Db      248  --IYQVAGGAGGAGYFAFAKSLNESRPNEMKYYWIGVDRDQEAEBGKYISKOGKEBNFLVY 305
Qy      307  SILKNLGYSVRSVYSLDLYTKKSNRNLAGFEFGKKSATVYIIGIDRNVYDIADTSLSEGNDK 366
Db      306  STLKQVGTYYVDI-----SNKAKECFPGGQ---VIYVSLKDKGVDLAVTNLSEBGK 354
Qy      367  KLAETEAISAEKKEFEKTKTIP 388
Db      355  K-----AVEDAKAKIILGGSVKVP 372

```

```

RESULT      5
D96986
probable lipoprotein, Med/BMP family [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: D96986
R:Noiling, J.; Britton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D96986
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-357 <RUP>
A:Cross-references: GB:AE001437; PIDN:AAK78679.1; PID:G15023581; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0702

```

|                       |                  |                   |            |             |
|-----------------------|------------------|-------------------|------------|-------------|
| Query Match           | 8.9%;            | Score 192;        | DB 2;      | Length 357; |
| Best Local Similarity | 23.3%;           | Pred. No. 0.0001; |            |             |
| Matches 94;           | Conservative 67; | Mismatches 157;   | Indels 86; | Gaps 18     |

```

01 20 ITDYSKISGLVNERKSEIEMAA--KADANKHGLMALVTAAGVJUNDNSFQSGEALIOOL 77
02 10 LITWIAVAGLEPAGGSSSTSSSGNSGSKDPRK--YKVGLSTDEGGJDNDSFQAGDEGICKA 66
03 78 CALUGGETTSDSSTA-ELECKYSSLANTKNKNWVLSGFOGDAF--RMLKIPENKOLFT 135
04 67 AKEXSVDPKALESKKKDDYOPILOSLIDNDSDLFEGVGYOMADLAVIARKYDPRKFAII 126
05 136 E-----KNIIIIIGIDWTDENVIPTGRVINTLYKTEEAGWLAGVYANASFLAKKFPSPD 188
06 127 DDAYDKQKKNL-----MSLVFKQDEGSLMG-----VIAGKM----- 158
07 189 TKRSAL-VIGGIGISPAVYDFIAGIAGIACAKANLKNKSUKKRIITTDKEIMLGGDVODTST 247
08 159 TKTNIGVGGGRDOPILNKFLFSLGIAGIAKTYN-----PUIYVK---NYTNDYSSTSK 208
09 248 -KERLEQASDKDPESTLLAIVAGPTEIFESIIANQNDR-----YLIGVDTQDSLVYTTKN 302
10 209 GKEVATSLYNGCIDIYVHAAGAGIGAF-DVAKELRDQGDQVMAIVGDKNQAGLPRYAD 267
11 303 KEFTSILKNLIGVSPSVSLDLYTKRSNRSNRLAGEPEPKKASTYVLLGIKDKRFVIDADTSL 362
12 268 VILTSMAVRVIDIATYVTKDLVKGKK-----FEGGVVES--FGIKEGGVGAPPS-- 315

```

|    |     |                       |                  |               |     |
|----|-----|-----------------------|------------------|---------------|-----|
| OY | 363 | GNDKRLATPAISEAK       | -----            | EEFEKRTIPAEVR | 393 |
|    |     | : : : : :             |                  | : : : :       |     |
| Db | 316 | --NKHVSEVLSLVDKYKKAII | DGKIVPDPVDKQOTFR | TDQIK         | 357 |

| RESULT | 6 |
|--------|---|
| H71340 |   |

C:Species: *Treponeema pallidum* subsp. *pallidum* (syphilis spirochete)  
C:Date: 07-Aug-1998 #sequence:revision 07-Aug-1998 #text\_change 23-Jul-1999

Riiser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Khaliq, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

A:Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770

A; status: nucleic acid sequence not shown; translation not shown  
A; molecule type: DNA

A: Experimental source: strain Nichols  
B: Chemical: T M, was from David H. O. T. was from  
C: Cross-references: GR:AE001211; GR:AE000520; NID:g33322582; PIDN:AAC63502.1; PID:g33322582

direct. Immun. 59, 3536-3546, 1991  
 A/T/tile: Characterization of the 35-kilodalton *Treponema pallidum* subsp. *pallidum* rec  
 A/Reference number: M25E05. UNTD.01372065

A: Accession: A43595  
A: Molecule type: DNA  
A: Residues: 1-10 / A' 12-158 / B' 160-353 (SCH)

Accession: AF037030; MIM:240030; FUN:CM0030.1; FID:350005  
 A>Note: this protein is shown to incorporate palmitic acid  
 C:Genetics:

ntGene: uniprot/ncbi  
 A: Start codon: GTG  
 C: Superfamily: ABC transporter vufn

F-21-353/Product: membrane lipoprotein tmnc #status predicted <MAT>  
F-1-20/Domains: signal sequence #status predicted <SIG>  
c/n/cn/cn/cn: located amino end /apparently initiating protein unobserved zone

F.21/Binding site: sn-2,3-diacylglycerol (Cys) (covalent) #status predicted

|                       |        |                    |       |             |
|-----------------------|--------|--------------------|-------|-------------|
| Query Match           | 8.4%;  | Score 181;         | DB 1; | Length 353; |
| Best Local Similarity | 23.7%; | Pred. No. 0.00049; |       |             |

52 MAIVTAGTVDNSFNQSGWEIQQQLGALTGGEITSYDSS\*-AELEGKYSSLANTNNK 110

Db 42 VGNVTTDGGDIDDKSPNQVWEGISRPAQENNAKCKYVTASTDAEYVPSLSAFADENMGVL 101

```

-      | | : | | : | : | : | :
102 VACSGFLVEA-----VIETSAFPPKKFLVIDAVQDRDNV-----VSAVFQNGESGF 149
Db

```

```
QY      171 LAGYANNSFLAKKFPSPDKRKRAI--VIGG--GISPAVTDFIAGIYLAGIAMNKLNSDK 226
```

227 TKITTDKIEINCFDVOdT-S\*EKERLEQIASK--DKPSTLLAVAG---PLTEIESDII 278

Db 194 -----PDIQVVEVANTFSDPQKGLAAKLYDSSVNVLEQVAGSTGNGVKEARDRR 246

Db 247 LINGDWNWICGVRDQYMDGVYDGSKSVLTSWVKRA---DYAAERISKMAVDGSPFG-300

[illegible]

QY 397 EIP 399 .1

Db 343 vfp 345

## RESULT 7

ABC transporter (lipoprotein) homolog yufN - Bacillus subtilis  
C:Species: Bacillus subtilis

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000

C:Accession: C70009

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C: Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Ch

A: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizal, A.; Gall

A: lech, J.; Harwood, C.R.; Henuat, A.; Hildbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

A: Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardin

A: Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

V. M.; Ogawa, K.; Ogilawa, A.; Oudaga, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet

A: ger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanlon,

A: lhorst, M.; Schleich, S.; Schotter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser

A: kuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A: Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A6580; MUID:98044033

A:Accession: C70009

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-350 &lt;KUN&gt;

A:Cross-references: GB:259120; GB:AL009126; NID:92635613; PIDN:CAB15143.1; PID:92635650

A:Experimental source: strain 168

C:Genetics:

A:Gene: yufN

C:Superfamily: ABC transporter yufN

Query Match 8.4%; Score 180.5; DB 1; Length 350;  
Best Local Similarity 23.9%; Pred. No. 0.00052;  
Matches 101; Conservative 60; Mismatches 144; Indels 85; Gaps 20;

37 IMAAANKKFKG-----LMAIVTAGTVNDNSFNQSGWEAIOQLGALG-----G 83  
 10 ILGAGNSKSEKSGSEKKNKRVANVTGVDGKSPNOSANEGLQAKGKGLKKKNG 69  
 84 EITVSSTAELEGRYSLSANTKNVWVLSGFQHDATFRLKIPENKOLFTEKNIIIG 143  
 70 YDLQSKSDADVTYTNMLKARENFDLIYGVYLMEDSIS---EIAQR-----KMTNFAI 121  
 144 ID-WDTENVIPDTGRYINLTYTEAGMLAGYANASFLAKRPPSPTRSAI-VIGGGIS 201  
 122 IDAVVDKDNV-----ASITFEKQSGSFLVGYAAL-----SSKSGKIGFVGCMES 166  
 202 PAVTFDIFAGYLAGIKAMVLKNSDKTKITTDKIEINLGFVDVT--SPKERLEQJASKDK 259  
 167 ELIKFEFVFRAGVQAVNPK-----AYVEVKYAGFDKADVKATRESH---YKSG 214  
 260 PSTLLAVAGPL-TEIFSDIIAN-----QNDRYLIGVDTDQ---SLVYTKTKNKEFTSL 309  
 215 VDVYHSAGAGTGVETE-AKMLKEDPKRDVWVIGVDKDYAEQVGSTDNVTLTSMV 273  
 310 KNLGVSVSVDLYTKKSNRNLAGFEFGKSAIVYIGIDRPVDIADTSLGNDKILA 369  
 274 K---KVTYEDV--TKRASDCKPFGG---TLTYGLDQDGVGISPKONLSDDVI- 321  
 370 TEAISEAKKEF-----EKKTKTPAE 390  
 322 -KAVDKKKKIIIDGLEIPATEKEKTFKAE 350

RESULT 8  
 F72418  
 basic membrane protein - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C:Accession: F72418

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic

C: Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,

C: M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome

A:Reference number: A7200; MUID:99287316

A:Accession: F72418

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-359 &lt;ARR&gt;

A:Cross-references: GB:AE001696; GB:AE000512; NID:9490582; PIDN:AD35196.1; PID:9498

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM0102

C:Superfamily: ABC transporter yufN

Query Match 8.3%; Score 178.5; DB 2; Length 359;  
Best Local Similarity 23.9%; Pred. No. 0.00071;  
Matches 94; Conservative 59; Mismatches 177; Indels 63; Gaps 16;

33 KSEFMKAADANKKFKG-----LMAIVTAGTVNDNSFNQSGWEAIOQLGALG-----S 91  
 2 RFEIVISLMIFVAFLEFKRVIVTVGGLGDKSFDGTWAGIKQAEELGIEAKVIOSE 61  
 92 TAELEGRYSLSANTKNVWVLSGFQHDATFRLKIPENKOLFTEKNIIIGDWTDE- 150  
 DB 62 QSDYIPNLSKAEEADLVFVAVGFMNTNDFKAKQYPT-----YVGIDITTPEG 112  
 151 NVIPTRYINLTYTEAGMLAGYANASFLAKRPPSPTRSAI-VIGGGIS-PATDEFA 209  
 DB 113 QILP--NVLTFTFEKQEAFLVGYAAL-----TKGMVGFQIPPIPERRY 161  
 210 GYLACIKAMNLSKDKTKI---TTDKIEINLGFVDVTSTKELEI----- 254  
 DB 162 GEAEIKITYSVLHK-KNKILNGTYODEPDKKGLDMSQFAEGADIVFNASGAGNGV 220  
 255 --ASKRKPSTLLAVAGPLTEIFSDIIANQNDRYLIGVDTDQSLVYTKKKNKFTSL 312  
 DB 221 IEMAKEKSA-LAGSDKLVLDIYTTNGKGFAGVMDODY---APGAVLSAMKRV 276  
 313 GYSVESVSDLYTKKSNRNLAGFEFGKSAIVYIGIDRPVDIADTSLGNDKILATEA 372  
 DB 277 DVASYGYVWAVE-----CTFEGRH---VLGISEDANVI--SPMKYTKGLVPRV 322  
 373 ISE---AKKEPEKTKTP-AEVKTKTEIPEM 401  
 DB 323 IAEILYLERLKMKEGTLKVPETQEEIDAEFVPI 355

RESULT 9  
 E70147  
 basic membrane lipoprotein B (bmbp) - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 29-Sep-1999  
 C:Accession: E70147; I40290; I40242  
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
 son, D.; Peterson, J.; Kertavage, A.R.; Quackenbush, J.; Salberg, S.; Hanson, M.; Yu  
 son, B.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
 A:Reference number: A70100; MUID:98065943  
 A:Accession: E70147  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-341 <KLE>  
 A:Cross-references: GB:AE001143; GB:AE000783; NID:92688279; PIDN:AC66758.1; PID:9268  
 A:Experimental source: strain B31  
 R:Simpson, W.J.; Cleplak, W.  
 FEMS Microbiol. Lett. 119, 381-388, 1994











GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:30:19 ; Search time 17.51 Seconds

(without alignments)  
935.372 Million cell updates/sec

Title: US-09-676-249a-4

Perfect score: 2155  
Sequence: 1 MMDKRTKREKRSADNKNQI.....KQPDKQI:SLDKITIDINNL 423

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 183   | 8.5         | 325    | 1  | BMPA_BORGA  |
| 2          | 181   | 8.4         | 353    | 1  | THRC_TREPA  |
| 3          | 180.5 | 8.4         | 350    | 1  | YUFN_BACSU  |
| 4          | 178   | 8.3         | 341    | 1  | BMPB_BORBU  |
| 5          | 173   | 8.0         | 357    | 1  | TCSA_LISIN  |
| 6          | 169   | 7.8         | 341    | 1  | BMPB_BORGA  |
| 7          | 168   | 7.8         | 357    | 1  | TCSA_LISMO  |
| 8          | 166   | 7.7         | 339    | 1  | BMPA_BORAR  |
| 9          | 165.5 | 7.7         | 341    | 1  | BMPD_BORBU  |
| 10         | 161   | 7.5         | 339    | 1  | BMPB_BORAR  |
| 11         | 153   | 7.1         | 341    | 1  | BMPB_BORAR  |
| 12         | 152.5 | 7.1         | 591    | 1  | Y040_MYCGE  |
| 13         | 146   | 6.8         | 353    | 1  | BMPC_BORBU  |
| 14         | 136   | 6.3         | 547    | 1  | MEBA_STRAU  |
| 15         | 129   | 6.0         | 498    | 1  | PEOB_STRPY  |
| 16         | 128   | 5.9         | 657    | 1  | Y040_MYCPN  |
| 17         | 124.5 | 5.8         | 981    | 1  | SCA4_RICFE  |
| 18         | 123.5 | 5.7         | 626    | 1  | CEJA_BCOLI  |
| 19         | 118.5 | 5.5         | 1140   | 1  | YD76_MYCPN  |
| 20         | 117   | 5.4         | 798    | 1  | Y000_MYCPN  |
| 21         | 114.5 | 5.3         | 848    | 1  | Y047_SCHPO  |
| 22         | 114.5 | 5.3         | 1290   | 1  | XCPG_XENLA  |
| 23         | 114   | 5.3         | 626    | 1  | PARC_BORBU  |
| 24         | 114   | 5.3         | 943    | 1  | LBPA_NEIMB  |
| 25         | 113.5 | 5.3         | 552    | 1  | FLCK_SALTY  |
| 26         | 113.5 | 5.3         | 998    | 1  | SCA4_RICAK  |
| 27         | 112   | 5.2         | 550    | 1  | FLIC_SHIFT  |
| 28         | 111.5 | 5.2         | 1111   | 1  | KIP1_YEAST  |
| 29         | 110.5 | 5.1         | 705    | 1  | PQCE_FLAME  |
| 30         | 110.5 | 5.1         | 1164   | 1  | BAG_STRAG   |
| 31         | 109.5 | 5.1         | 672    | 1  | Y0A2_MYCPN  |
| 32         | 109   | 5.0         | 445    | 1  | MAPA_STRMU  |
| 33         | 108.5 | 5.0         | 626    | 1  | CEIB_BCOLI  |

|    |       |     |      |   |            |                    |
|----|-------|-----|------|---|------------|--------------------|
| 34 | 108.5 | 5.0 | 944  | 1 | LBPA_NEIMA | O91k4 neisseria m  |
| 35 | 108   | 5.0 | 3418 | 1 | BRC2_HUMAN | P51587 homo sapien |
| 36 | 107.5 | 5.0 | 794  | 1 | YC84_MYCPN | P75493 mycoplasma  |
| 37 | 107.5 | 5.0 | 2022 | 1 | ANT1_ONCVO | P21249 onchocerca  |
| 38 | 107   | 5.0 | 770  | 1 | ECBI_HUMAN | P42892 homo sapien |
| 39 | 107   | 5.0 | 1287 | 1 | RPO1_FOWPV | O91593 fowlpox vir |
| 40 | 106   | 4.9 | 429  | 1 | YN8H_YEAST | P53729 saccharomyc |
| 41 | 106   | 4.9 | 624  | 1 | HTPG_BUCAI | P57555 buchnera ap |
| 42 | 106   | 4.9 | 1076 | 1 | RPOB_ASTLO | P27059 astasia lon |
| 43 | 105.5 | 4.9 | 1176 | 1 | SLAP_BACST | P38537 bacillus sp |
| 44 | 105.5 | 4.9 | 1178 | 1 | YN17_YEAST | P48231 saccharomyc |
| 45 | 105   | 4.9 | 317  | 1 | PLC_LISMO  | P34024 listeria mo |

## ALIGNMENTS

| RESULT ID | 1  | BMPA_BORGA | STANDARD; | PRT; | 325 AA.                        |
|-----------|--|------------|-----------|------|--------------------------------|
| AC        | 031357; 031360;  |            |           |      |                                |
| DT        | 15-DEC-1998 (Rel. 37, Last sequence update)  |            |           |      |                                |
| DT        | 15-DEC-1998 (Rel. 37, Last sequence update)  |            |           |      |                                |
| DE        | Basic membrane protein A precursor (Immunodominant antigen P39) (Fragment).  |            |           |      |                                |
| GN        | BMPA.  |            |           |      |                                |
| OS        | Borrelia garinii.  |            |           |      |                                |
| OC        | Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.   |            |           |      |                                |
| OX        | NCBI_Taxid=29519;  |            |           |      |                                |
| RN        | (1)  |            |           |      |                                |
| RP        | SEQUENCE FROM N.A.   |            |           |      |                                |
| RC        | STRAIN-PBI, AND PLI;   |            |           |      |                                |
| RX        | MEDLINE=98010210; PubMed=9350727;  |            |           |      |                                |
| RA        | Roesler D., Hauser U., Wilske B.;  |            |           |      |                                |
| RT        | "Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnosis."   |            |           |      |                                |
| RL        | J. Clin. Microbiol. 35:2752-2758(1997).  |            |           |      |                                |
| CC        | -1- FUNCTION: NOT KNOWN: IMMUNOGENIC PROTEIN.  |            |           |      |                                |
| CC        | -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).   |            |           |      |                                |
| CC        | -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.   |            |           |      |                                |
| CC        | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ). |            |           |      |                                |
| DR        | EMBL; X87244; CAA55883.1; -  |            |           |      |                                |
| DR        | EMBL; X87238; CAA55877.1; -  |            |           |      |                                |
| DR        | InterPro; IPR003760; Bmp.  |            |           |      |                                |
| DR        | InterPro; IPR000437; Prok_Lipoprot.  |            |           |      |                                |
| DR        | Pfam; PF02608; Bmp; 1.   |            |           |      |                                |
| DR        | PROSITE; PS00013; PROKAR_LIPOPROTEIN; PARTIAL.   |            |           |      |                                |
| KW        | Antigen; Membrane; Lipoprotein; Signal.  |            |           |      |                                |
| FT        | NON_TER  | 1          |           |      |                                |
| FT        | SIGNAL   | <1         | 3         |      | PROBABLE.                      |
| FT        | CHAIN  | 4          | 325       |      | BASIC MEMBRANE PROTEIN A.      |
| FT        | LIPID  | 4          | 4         |      | N-ACYL DIGLYCERIDE (PROBABLE). |
| FT        | VARIANT  | 8          | 8         |      | G -> D (IN STRAIN PLI).        |
| FT        | VARIANT  | 12         | 12        |      | S -> N (IN STRAIN PLI).        |
| FT        | VARIANT  | 93         | 93        |      | P -> S (IN STRAIN PLI).        |
| FT        | VARIANT  | 111        | 111       |      | T -> A (IN STRAIN PLI).        |
| FT        | VARIANT  | 180        | 180       |      | N -> D (IN STRAIN PLI).        |
| FT        | VARIANT  | 205        | 205       |      | S -> G (IN STRAIN PLI).        |
| FT        | VARIANT  | 239        | 239       |      | I -> V (IN STRAIN PLI).        |
| FT        | VARIANT  | 243        | 243       |      | A -> S (IN STRAIN PLI).        |
| FT        | VARIANT  | 250        | 250       |      | S -> A (IN STRAIN PLI).        |
| FT        | VARIANT  | 257        | 257       |      | N -> S (IN STRAIN PLI).        |



RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RA Oudega B., Koningssteijn G., Duesterhoeft A., Hilbert H.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
 CC (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: 293937; CAB07936.1; -  
 DR EMBL: 299120; CAB5143.1; -  
 DR Subtilist; BG12349; yuFn.  
 DR InterPro: IPR003760; Bmp.  
 DR Pfam: PF02608; Bmp; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 DR Hypothetical protein; Membrane; Lipoprotein; Signal;  
 DR Complete proteome.  
 FT SIGNAL 1 13 PROBABLE.  
 FT CHAIN 14 350 HYPOTHEITICAL LIPOPROTEIN YUFN.  
 FT LIPID 14 14 N-ACYL DIGLYCERIDE (PROBABLE).  
 SO SEQUENCE 350 AA; 37349 MW; 16D5176A52A99284 CRC64;

Query Match 8.4%; Score 180.5; DB 1; Length 350;  
 Best Local Similarity 25.9%; Pred. No. 0.00027;  
 Matches 101; Conservative 60; Mismatches 144; Indels 85; Gaps 20;

QY 37 IMARKADANKHFG-----LNMAITYAGCTVNDNSFNOSGWEAIOQLALTC-----G 83  
 DB 10 ILGACNSEKSSGSGCKNFSAAMVTGCVGVDKSFNSMGWICQAFCKENILKGRNG 69  
 QY 84 EITSVDSSTAELEKXSSLANTRKNWVLSGFQGDFAFTWMLKIPERKQFTENITLG 143  
 DB 70 YDVLQSKSDADYTNLTKLARENFDLTYGVGLMEDSIS---ELADNR-----KNTFAI 121  
 QY 144 ID-WTDENVIPTRGYINLTYTEAGWLAGYNASFLAKKFPSPDKNSAI-VIGGIS 201  
 DB 122 IDAVVVKDNY-----ASTIFKEQSGSLVGVAAAL-----SKSGKIGFVGGMES 166  
 QY 202 PATYDPIAGLAGIKAMNKNKSDKRTITDKLEINLGFVQDT--STKERLQIASKDK 259  
 DB 167 ELTKKEVGRAGVQAVNFK-----AVEEVKVGGEFKADVGKATVATESM---YKSG 214  
 QY 260 PSTLLAVAGPL-TEIFSDILAN-----QNDRYLIGVDIDQ--SLVYTKRKKFTSIL 309  
 DB 215 VDVYHSAGATGTVFTE-AKNLKEDPKRDVAVIGVDKQVGEQVGGEDDVTILTSMV 273  
 QY 310 KNLGYVSFVSLDYTKRKSNNLAGFEFGKSAVYLGITKRPVDIADTSLSGNDKKLA 369  
 DB 274 K-----KVDYVEVDY-TKKASDGKFPGE-----TLTYGDODGVGISPSKQMLSDVI- 321  
 QY 370 TEAISEAKKEF-----EKKTIIPAE 390  
 DB 322 -KAVDKMKKKIIDGLEIPATEKELTKFKAE 350

RESULT 4  
 BMBP.BORBU STANDARD: PR: 341 AA.  
 AC Q45011; Q07954; Q31317; Q50168; Q44856;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Basic membrane protein B precursor.  
 GN BMBP OR BR0382.  
 OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 ON NCBI\_TaxID-139;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SH-2-82;  
 RX MEDLINE-94327086; PubMed-8050720;  
 RA Simpson W.J., Cieplak W., Schrupf M.E., Barbour A.G., Schwan T.G.;  
 RT Nucleotide sequence and analysis of the gene in Borrelia burgdorferi  
 RL encoding the immunogenic P39 antigen.";  
 RT FEMS Microbiol. Lett. 119:381-388(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-297;  
 RX MEDLINE-97132632; PubMed-8978084;  
 RA Aron L., Toth C., Godfrey H.P., Cabello F.C.;  
 RT Identification and mapping of a chromosomal gene cluster of Borrelia  
 RL burgdorferi containing genes expressed in vivo.";  
 RN FEMS Microbiol. Lett. 145:309-314(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 35210 / B31;  
 RX MEDLINE-98010210; PubMed-9350727;  
 RA Roessler D., Hauser U., Wilske B.;  
 RT Heterogeneity of BmpA (P39) among European isolates of Borrelia  
 RL burgdorferi sensu lato and influence of interspecies variability on  
 RN serodiagnosis.";  
 RN J. Clin. Microbiol. 35:2752-2758(1997).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 35210 / B31;  
 RX MEDLINE-98065943; PubMed-9403685;  
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,  
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Woldman J.,  
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
 RA Uitterlidge T., Wathey L., McDonald L., Artiach P., Bowman C.,  
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 RA Smith H.O., Venter J.C.;  
 RT Genomic sequence of a Lyme disease spirochete, Borrelia  
 RL burgdorferi.";  
 RL Nature 390:580-586(1997).  
 RN [5]  
 RP SEQUENCE OF 1-179 FROM N.A.  
 RC STRAIN-212;  
 RX MEDLINE-95111614; PubMed-7812434;  
 RA Ojaimi C., Davidson B.E., Saint-Girons I., Old I.G.;  
 RT Conservation of gene arrangement and an unusual organization of rRNA  
 RT genes in the linear chromosomes of the Lyme disease spirochaetes  
 RL Borrelia burgdorferi, B. garinii and B. afzelii.";  
 RL Microbiology 140:2931-2940(1994).  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
 CC (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L24194; AAA72407.1; -  
 DR EMBL: U49938; AAC44713.1; -  
 DR EMBL: X81517; CAA57227.1; -  
 DR EMBL: AE001143; AAC66758.1; -  
 DR EMBL: L35050; AAC41402.1; -  
 DR TIGR: BR0382; -  
 DR InterPro: IPR003760; Bmp.  
 DR Pfam: PF02608; Bmp; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 DR Membrane; Lipoprotein; Signal; Complete proteome.  
 KW

FT SIGNAL 1 14 PROBABLE.  
 FT CHAIN 15 341 BASIC MEMBRANE PROTEIN B.  
 FT LIPID 15 341 N-ACYL DIGLYCERIDE (PROBABLE).  
 FT VARIANT 45 45 S -> A (IN STRAIN 212).  
 FT VARIANT 233 233 A -> T (IN STRAIN B31).  
 FT VARIANT 318 318 V -> I (IN STRAIN B31).  
 FT CONFLICT 53 53 R -> P (IN REF. 3).  
 FT CONFLICT 67 67 S -> P (IN REF. 3).  
 SQ SEQUENCE 341 AA; 37549 MW; C26A0D4B1D52P39F CRC64;

Query Match 8.3%; Score 178; DB 1; Length 341;  
 Best Local Similarity 26.9%; Pred. No. 0.0037;  
 Matches 101; Conservative 58; Mismatches 124; Indels 92; Gaps 25;

OY 44 ANKHGFLNMAIYAGTGVNDNSFNOSGWEAIOALGALGCEITSVDSVAELEGKYS-- 101  
 DB 24 SSKKIKISMLV---DGVLDKSPFNSSANALRLKDKPFENIEEVS--CAISGVYSSIV 78  
 DB 102 --LAINTNKN---VWVLSGFHGDFAETRLKIPENKQLETERNII--ILGIDMTDTENVY 153  
 DB 79 SLDNLKRRGSDLIW-LVGVMLTDA--SLVASENPKI--SYGIIDPIYGDVQIPEMLI 133  
 OY 154 PGRGINTLYKTEEGAGMLAGY--ANASFLAKKFPSPDPTKRSALVIG--GGISPAVTD-FI 208  
 DB 134 A-----VVERVEQGAFLAGYIAAKKSFSGK-----IGFIGMGKGNIVDAER 174  
 OY 209 AGYLAGIKAMNLKNSDKKRTITDKIEINLGFVDVDTSTKERLEQIASK--DKPSTLLAV 266  
 DB 175 YYESGAKKAN-KDIEIIEISYNSFSVDVIG-----RTIASMYKSGIDVYIF 221  
 OY 267 AGPLREI-FSDIIANQND-RYLIGVTDQSLVYTKRNFETSIKNLGYSVFVSLDLY 324  
 DB 222 AAGLAGIGVIEAKNLDGQYVIGADQDS--YLAPEKN-FITSVIKNICDALYLTIGE-Y 277  
 OY 325 TKKSNRLNLAGFEFGKSKATYVVLGIDKRVVDIADTSLEGNCKKATEAISEAKKEEET 384  
 DB 278 INNNAV-----WEGK-----VVGMLRGDVGILPMAN-----EF-EYI 310  
 OY 385 KTIPEAEVKTLEIP 399  
 DB 311 KYLERKIVKKEIIVP 325

RESULT 5  
 TCSA-LISIN STANDARD; PRT: 357 AA.  
 092BM7:  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE CD4+ T cell-stimulating antigen precursor.  
 GN TCSA OR LIN1425.  
 OS *Listeria innocua*.  
 OC Bacteria; Filmylcoetes; Bacilllus/Clostridium group;  
 OC Bacillus/Staphylococcus group; *Listeria*.  
 OX NCBI\_TaxID=1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIP 11262 / Serovar 6a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Franseul L., Buchrieser C., Rusnlok C., Amend A.,  
 Baquero P., Berche P., Blocher H., Brandt P., Chakraborty T.,  
 Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,  
 Domant E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 Entian K.-D., Febl H., Garcia-del Portillo F., Garrido P.,  
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kuzepak G.,  
 Madueno E., Maitouram A., Mata Vicente J., Ng E., Nedjari H.,  
 Nordstok G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 Rammel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
 Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.,  
 RT Comparative genomics of *Listeria* species.;

RL Science 294:849-852(2001).  
 CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
 CC (Probable).  
 CC -I- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AL596168; CAC96656.1; -  
 DR Listlist: L1N01425; -  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Antigen; Membrane; Lipoprotein; Signal; Complete proteome.  
 FT SIGNAL 1 22 PROBABLE.  
 FT CHAIN 23 357 CD4+ T CELL-STIMULATING ANTIGEN.  
 FT LIPID 23 23 N-ACYL DIGLYCERIDE (PROBABLE).  
 SQ SEQUENCE 357 AA; 38357 MW; 2EDA2AA9820A3FBC CRC64;

Query Match 8.0%; Score 173; DB 1; Length 357;  
 Best Local Similarity 24.2%; Pred. No. 0.00079;  
 Matches 101; Conservative 61; Mismatches 138; Indels 118; Gaps 23;

OY 27 SGLV-----NEKSEIMAKADANKHFLNMAIYAGGVNDNSFNOSGWEAIOQLG 78  
 DB 16 SGVVLGACGSSDDKKS-----GDKSKDPTVAMVTDGTGVDNRSNOSAMEGLQKFG 69  
 OY 79 ALT-----GGEITSVDSSTAELEGKYSILANTKNVWLS-----GFOHGDFAETRLK 126  
 DB 70 KANDEKGTGVNYLQSASEADYK-----TMLNFAVSDVLIIGIYKIKLDAIEEYK 123  
 OY 127 -IPEKQLETERNIIILGIDMTDTENVYPTGRYINLTYKTEEGAGLAGYANASFLAKFP 185  
 DB 124 OKPKNOFALVDDTI-----DDRNVVSIK-----FRDNGSYLVGVAGL----- 163  
 OY 186 SDPTKRSALVIGGISPAVTD-FINGYLAGIKAMNLKNSDKKRTITTKIEINLGFVDO- 243  
 DB 164 --TTTKNKVGFYGVKGAVIDFEAGFTAGVKA-----VNPNAIDVQY 205  
 OY 244 --DTSTKERLEQIASKDKPS-----TLAVAGPLTEIFSDIIANQND-----YLDIGVT 291  
 DB 206 ANDPAKADKGGQIASSMVSQGVDFHAAAGTGVNCFPAE-AKNLKKRPSRAVWVIGDR 264  
 OY 292 DO-----SLVYTKRKNFTTSILKNLGYSVFVSLDLYTKKSNRLAG-FEFGKKSAT 344  
 DB 265 DQWDEGKYIANDGKDYNTLTSEIKRVDIIV---DDLAT-----RTKAGDFPGGTFK--- 312  
 OY 345 VYLGIKDRVVDIADTSLEGNCKKATEAISEAKKEEETKTIPEAEVKTLEIPEMP 402  
 DB 313 IYEGDKDAVGLSE-----HODNISKDVAKYE---ETKQKIVDGD-----IKVEKP 357

RESULT 6  
 BMPB\_BORCA STANDARD; PRT: 341 AA.  
 AC 031362;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Basic membrane protein B precursor.  
 GN BMPB.  
 OS *Borrelia garinii*.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.  
 OX NCBI\_TaxID=29519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PRT;  
 RX MEDLINE=96010210; PubMed=9350727;  
 RT Roessler D., Hauser U., Wilske B.;



\*Heterogeneity of Bmpa (P39) among European isolates of *Borrelia burgdorferi* sensu lato and influence of interspecies variability on serodiagnosis.";  
 RL J. Clin. Microbiol. 35:2752-2758(1997).  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X81518; CA57238.1; -  
 DR InterPro: IPR003760; Bmp.  
 DR Pfam: PF02608; Bmp; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Membrane; Lipoprotein; Signal.  
 FT SIGNAL 1 14 PROBABLE.  
 CHAIN 15 341 BASIC MEMBRANE PROTEIN B.  
 LIPID 15 15 N-ACYL DIGLYCERIDE (PROBABLE).  
 SQ SEQUENCE 341 AA; 37236 MW; 442BEF0BEDDC9A CRC64;  
 Query Match 7.8%; Score 169; DB 1; Length 341;  
 Best Local Similarity 26.8%; Pred. No. 0.0013;  
 Matches 91; Conservative 55; Mismatches 117; Indels 76; Gaps 20;  
 QY 59 GIVNDNSFNOSGWEAIOQLALGTGELTSDVSTAELEGKYS--LANTNKN---VW 110  
 DB 36 GVLDDKSFNSANRALLRLLEDPEPENIEKVESMA--SGVYSVSDLDLKNKMGSDLIW 93  
 QY 111 VLSGFQHGDAFTWMLKPEKOLETEKNITILG-IDMTDENVYIPTRYINLYTKTEAG 169  
 DB 94 -LVGYMLTDA-----SLVSLNPKLSYGIIDPVSDDOVQIPKNLGVVERIRIOGA 143  
 QY 170 WLAGYANASFLAKFPSPDPTKRSATVIGGISPAVTD-FLAGIAGIKAMNLSKDKTKT 228  
 DB 144 FLAGY-----IAAKKVS-----GKIGFVGKGDYDAFRTYAG-----AKADKIE 189  
 QY 229 ITTDR---TEINLGFVDVDTSTKERLEQIASK--DKPSTLLAVAGHLTEIFSDIANON 282  
 DB 190 IVSEYSNFSFDVNI-----RAIANKMYAKGIDIIHFAAGLAGI--GVIEAPK 235  
 QY 283 D-----RLIVDVDTSLVTKTKNKFSTILKNGSVESVLSDLTYKKSNSRLAGFER 338  
 DB 236 ELGDGYVIGADDDSHL---APRNFITSYIKKVGDAIYLTSE-VYKNNNT-----WEG 286  
 QY 339 GKKSATVYLGIDKDFDIADTS---LEGNDKILATEAI 373  
 DB 287 GK---IIOMGLRDGVVGLSNANKFEYIKVIERKIVNEEI 322  
 RESULT 7  
 TCSEA.LISMO STANDARD; PRT; 357 AA.  
 AC 048754;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE CD4+ T cell-stimulating antigen precursor.  
 GN TCSEA OR LMO1388.  
 OS *Listeria monocytogenes*.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; *Listeria*.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BGD-e / Serovar 1/2a;  
 RX MEDLINE=21537279; PubMed=11679669;

RA Glaser P., Frangoul L., Buchrieser C., Rusnick C., Amend A.,  
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Cheretouani F., Couve E., de Darvar A., Dehoux P.,  
 RA Dommann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 RA Ertan K.-D., Fsih H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,  
 RA Madueno E., Maitounan A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.,  
 RT Comparative genomics of *Listeria* species.";  
 RL Science 294:849-852(2001).  
 RN [2]  
 RP SEQUENCE OF 1-252 FROM N.A.  
 RC STRAIN-BSE0-1167;  
 RX MEDLINE=96096448; PubMed=7500019;  
 RA Sanderson S., Campbell D.J., Shastri N.;  
 RT Identification of a CD4+ T cell-stimulating antigen of pathogenic  
 RT bacteria by expression cloning.";  
 RL J. Exp. Med. 182:1751-1757(1995).  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: AL591979; CAC9466.1; -  
 DR EMBL: S80336; AAB35725.2; ALT-TERM.  
 DR LISTLOC: LMO01388;  
 DR InterPro: IPR003760; Bmp.  
 DR Pfam: PF02608; Bmp; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Antigen; Membrane; Lipoprotein; Signal; Complete proteome.  
 FT SIGNAL 1 22 PROBABLE.  
 CHAIN 23 357 CD4+ T CELL-STIMULATING ANTIGEN.  
 LIPID 23 23 N-ACYL DIGLYCERIDE (PROBABLE).  
 SQ SEQUENCE 357 AA; 38415 MW; 83605B8B6419C8D1 CRC64;  
 Query Match 7.8%; Score 168; DB 1; Length 357;  
 Best Local Similarity 23.8%; Pred. No. 0.0016;  
 Matches 98; Conservative 62; Mismatches 145; Indels 106; Gaps 22;  
 QY 26 ISGLVNERKSEIMAKADANKHFLNMAIYTAGTVNDNSFNOSGWEAIOQLALT---- 81  
 DB 19 ILGACSSSDPKSKSDKSKDF--TYAMVTDIGVDKRSFNSAMGLOKFGKANDMEK 76  
 QY 82 GEITVSDSTAELEGKYSLSIANTNNKVVVLS-----GFQHGDAFTWMLK-IPENKQ 132  
 DB 77 GTDGYNYLOSSEADYK-----TNLNTAVRSDYDLTYGKRLKDAIEVSKQPKNOF 130  
 QY 133 LFTFKNIIILGIDMTDENVYIPTRYINLYTKTEAGVLAGYANASFLAKFPSPDKRS 192  
 DB 131 AIVDDTI-----DDRDNVYSIG-----FKNDGSGYLVGVAGL-----TTKTN 168  
 QY 193 AIVGGISPAVTD-FLAGIAGIKAMNLSKDKTKITTDKIEINLGFVDQ--DTSTK 248  
 DB 169 KVGFGVKGKVIDRFNAGFTAGYKA-----VAPNAQIDVQVANDPRAKA 212  
 QY 249 ERLEQIASDKPS---TLAVAGLTFEISDIANONDR-----YLGVDTQ----- 293  
 DB 213 DKGOQIASMYSSGVDTYFHAAGTGNGVFAE-AKNLKKDPSRAVWVIGVDRDQWDEK 271  
 QY 294 -SLVYTKTKNKFSTILKNGSVFVLSDLTYKKSNSRLAG-FEFGKKSATVYLGIDK 351  
 DB 272 VTANDKDVNVTLTSEIKRDIADV-----EDLAT-----RAKAGDFPGTK--IEYGLDK 319

```

OY 352 REVADIADTSLEGNDKKLATEAISEAKKEFEKTKTIPAEVRKLTIEPMP 402
DB 320 DAVGISE-----HODNISKVLAKVE---EYKQKIVDGD-----IKVPEKP 357

RESULT 8
BMPA_BORAF STANDARD: PRT: 339 AA.
AC O31280; O31281; O31282; O31283;
AD 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
GN Basic membrane protein A precursor (Immunodominant antigen P39).
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OY NCBI_TaxID=29518;

[1]
SEQUENCE FROM N.A.
RC STRAIN-PKO, PWUDI, PLE, AND PLJ7;
RX MEDLINE=98010210; PubMed=9350727;
RA Roessler D., Hauser U., Wilske B.;
RT "Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnostics."
RL J. Clin. Microbiol. 35:2752-2758(1997).
CC -1- FUNCTION: NOT KNOWN; IMMUNOGENIC PROTEIN.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).
CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X81516; CAA57236.1; -
DR EMBL; X97237; CAA65876.1; -
DR EMBL; X97239; CAA65878.1; -
DR EMBL; X97241; CAA65880.1; -
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp. 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC Antigen; Membrane; Lipoprotein; Signal.
FT SIGNAL 1 17 PROBABLE.
FT CHAIN 18 339 BASIC MEMBRANE PROTEIN A.
FT LIPID 18 18 N-ACYL DIGLYCERIDE (PROBABLE).
FT VARIANT 125 125 A -> S (IN STRAIN PLE).
FT VARIANT 214 214 I -> T (IN STRAIN PLJ7).
FT VARIANT 229 229 A -> P (IN STRAIN PKO).
FT VARIANT 254 254 I -> V (IN STRAIN PLJ7).
FT VARIANT 268 268 L -> F (IN STRAINS PLE AND PKO).
SQ SEQUENCE 339 AA: 36966 MW: 15648 Cys: 64
Query Match 7.7%: Score 166; DB 1; Length 339;
Best Local Similarity 23.3%; Pred. No. 0.002;
Matches 89; Conservative 68; Mismatches 143; Indels 82; Gaps 19;
OY 49 GLNMAIVTAGTVNDNSFQSGMEAIQQLGALTGGEITSVDST-----AELEGKYSILAN 104
DB 27 GIPKVSIVIDGTFDDKSNESALNGVKLKEFEIYLVKSSSTSYSLDLEG-----LKD 82
OY 105 TKNKNWVSGVOHQDAFTFRMLKIPENKQILFTEKNIIILCIDMTDTENVIPGRYINLTLYK 164
DB 83 AGSNILWIGIKFSDV-AKAVSLQNSEKKYAI-----IDPVYSNEPIP-ANLWGKTFR 133
OY 165 TEEAGWILAGYNASFLAKKPPSDPKRSAYIIGGISPAVMD-FLAGLAGIKAMNKLNS 223

```

```

DB 134 AOEGLAFGLYIAKV-----SKTGKIGFLGIEGDIDYDAFRYGYEAGAKVAN---- 180
OY 224 DKRTITTDKI-----EINLGEVDQDTSKERLEQJASKDKPSTLLAVGLPIREISDIIA 279
DB 181 -KDIKIFSGYIGSFSDLENG---RSVATKMYSDGI-----DIHHAAGLGICAIIEVA 229
OY 280 NO-NDRYLIGVDTPQSLVYTKTKKFFTSILKNGYVFSVSLDYTKKSNRLAAGFE 337
DB 230 KELGSGHYIIGVDEQSYL---APNNVITSTKDVGRSL-NLLTSNYLKTWT-----FE 279
OY 338 FKKKSATYVLGIKDRFVIDADTSLEGNDKKLATEAISEAKKEFEKTKTIPAEVRKLTLE 397
DB 280 GCK---LNYGLKEGVGV-----RNPMPKEFV-----EKEDLSKSIITKEVI 323
OY 398 IPEMPDKOPDKQOESLDKLTID 419
DB 324 VP-----YKKESEYKFLKE 337

RESULT 9
BMPD_BORBU STANDARD: PRT: 341 AA.
AC 044743;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Basic membrane protein D precursor.
GN BMPD OR BB0385.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OY NCBI_TaxID=139;

[1]
SEQUENCE FROM N.A.
RC STRAIN-JDI;
RX MEDLINE=96178617; PubMed=8606088;
RA Ramamoorthy R., Rovinelli L., Philipp M.T.;
RT "Molecular characterization, genomic arrangement, and expression of bmpD, a new member of the bmp class of genes encoding membrane proteins of Borrelia burgdorferi."
RL Infect. Immun. 64:1259-1264(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A., Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M., Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M., van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C., Uitterlind S., Fujii C., Cotton M.D., McDonald L., Artlich P., Bowman G., Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."
RL Nature 390:580-586(1997).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).
CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U35450; AAC43984.1; -
DR EMBL; AE001144; AAB91505.1; -
DR TIGR; BB0385; -
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp. 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.

```



Db 326 -SNKSEYEFLKE 337

## RESULT 11

EMBL\_BORAF STANDARD: PRT: 341 AA.

AC O31284;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Basic membrane protein B precursor.  
 GN MBP.  
 OS Borrelia afzelii.  
 CC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=29518;  
 [1]

SEQUENCE FROM N.A.

STRAIN-PRO; MEDLINE=98010210; PubMed=9350727;

RA Roessler D., Hauser U., Wilske B.;  
 "Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnostics";

RT J. Clin. Microbiol. 35:2752-2758(1997).  
 RL -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL; X81519; CA57239.1; -.

DR InterPro: IPR003760; Bmp.

DR Pfam: PF02608; Bmp; 1.

DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.

KW Membrane: Lipoprotein; Signal

FT SIGNAL 1 14 PROBABLE.

FT CHAIN 15 341 BASIC MEMBRANE PROTEIN B.

FT LIPID 15 15 N-ACYL DIGLYCERIDE (PROBABLE).

SEQUENCE 341 AA; 37198 MW; 188EA3E3D54ACDB1 CRC64;

Query Match 7.1%; Score 153; DB 1; Length 341;  
 Best Local Similarity 25.4%; Pred. No. 0.012;  
 Matches 90; Conservative 55; Mismatches 126; Indels 84; Gaps 19;

OY 59 GVNNSNSFSGWEALIOGLALTGGEITSVDSSTAELEGKYS-----LANTKN-----VW 110  
 DB 36 GYLDKRSNSNANELLRLKADFPENIEKVS--VSGYSSSYSDLDNLKMSNGLIW 93  
 OY 111 VLSGFQHD-AETRWLKIPEKQKFTKRNII--ILGIDMTDTEVPIPGRYINLTYSKTE 167  
 DB 94 -LVGYMLADVLSVLENDPE-----INVGIIDPIYGDVQIIPKNL-----IGVFRLEQ 141  
 OY 168 AGMLAGYANASFLAKKFSDDPKRSAYIYGGISPAVTD-FLAGLAGITKAMN-----L 220  
 DB 142 GAFLAGY-----IAAKKSVS-----SKIGFIGVGKGDIVDAFFGYEAGAKYANKGIEIYS 192  
 OY 221 KNSDKKTITTDKIEIN-----LGFDVODTSTKERLEQJASKRKPSTLLAVAGPLREISD 276  
 DB 193 EYSNFSIDILARVANKMYSGKIDIIHPAA--GLAGVGVIEAPKEL----- 237  
 OY 277 IIANONDRYLIGVDPDQSLVYTKTKNKEFTSLIKMLGYSVESVSLDLYTKKSNHLAGF 336  
 DB 238 ----GQYVYVIGADQDQSHLAPK---NFTSVYKKNVGDALYITSE-----SLKNDNW 284  
 OY 337 EFGKSAIVYIGIKRFDVADIADTSLGNDKLLATEAISEAKKEFEKTKTTPAE 391

Db 285 EGGK---IVQMGIRDGVVGLSNAN-----EPEYIKLERKINKKEIIVPCNQ 328

## RESULT 12

Y040\_MYCGE STANDARD: PRT: 591 AA.

AC P47286;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical lipoprotein MG040 precursor.  
 GN MG040.

OS Mycoplasma genitalium.  
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 CC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 [1]

SEQUENCE FROM N.A.

STRAIN-ATCC 33530 / G-37;  
 MEDLINE=96026346; PubMed=7569993;

RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,  
 Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 Tomb J.-F., Dougherty B.A., Boff K.F., Hu P.-C., Lucier T.S.,  
 Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 "The minimal gene complement of Mycoplasma genitalium";  
 Science 270:397-403(1995).

RL [2] SEQUENCE OF 448-517 FROM N.A.  
 RP STRAIN-ATCC 33530 / G-37;  
 RC MEDLINE=94075230; PubMed=8253680;  
 RA Peterson S.N., Hu P.-C., Boff K.F., Hutchison C.A. III;  
 "A survey of the Mycoplasma genitalium genome by using random sequencing";  
 J. Bacteriol. 175:7918-7930(1993).

RT J. Bacteriol. 175:7918-7930(1993).  
 RL -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Potential).  
 CC -1- SIMILARITY: SOME, TO T.PALLIDIUM TMC.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL; U39683; AAC71256.1; -.

DR EMBL; U02125; AAD12400.1; -.

DR TIGR; MG040; -.

DR InterPro: IPR003760; Bmp.

DR Pfam: PF02608; Bmp; 1.

DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.

KW Hypothetical protein: Lipoprotein; Membrane; signal;

FT SIGNAL 1 30 POTENTIAL.

FT CHAIN 31 591 HYPOTHETICAL LIPOPROTEIN MG040.

FT LIPID 31 31 N-ACYL DIGLYCERIDE (POTENTIAL).

SEQUENCE 591 AA; 64019 MW; 1A48FA31BB7E0928 CRC64;

Query Match 7.1%; Score 152.5; DB 1; Length 591;  
 Best Local Similarity 20.8%; Pred. No. 0.028;  
 Matches 99; Conservative 84; Mismatches 135; Indels 159; Gaps 27;

OY 44 ANKHGELMAIVTAG--GVNDNSFSGWEALIOGLALTG--EITSVDSSTAELEGK-- 98  
 DB 33 ATKIFDSSVQLLVSNFSLADKRSQMSREGIRSFKKSGVDLPADSSQLQ-EGNGL 91  
 OY 99 -----YSSLANTKNKYNVLSGROHGDAFTRWLKIPEKQKFTKRNIIIGI 144

```

Db 92 MKRPGTSLDRATENNINNGSDVIATGFNO-----QESLQATSDDIRF--- 138
Qy 145 DWITDENVIPTG-----RYINLYTEAGWLAGYANAS 178
Db 139 -GSDKSLAKTGFIFVDGAIEKFNKRNQVPOKSPPTNISAFRSDOSFLTGATAV 197
Qy 179 FL-----AKRPSDPTKRSAYIGG--GIS-PAYTDFIAGYAGIKAMN-----L 220
Db 198 YLNLNDQVFLDKSGWSTNNSNNELTVSGFVGIALPSTLSFLNGFRLGIAYFNEVIYKHL 257
Qy 221 KMS-DKKTITITDK-----IEINIG-----FDVOD-----TST 247
Db 258 SDHSDSSAOVTTSKQTVKQLOVANGKRIKTKWISPGQSGEITINIDHOSGSESDT 317
Qy 248 KERLEQIASK--DK-PSTLLAVAGPTE-IFSDIIANQNDRYLIGVTDQSL-----VYTK 299
Db 318 EPRATITANNLIDKGNVAILIPIAGPQTNLVVQIARQAHTAVIAGVDSAGQELLDINIDAP 377
Qy 300 TKNKF-----PSTLKNLGYSVFVSLDYLYTKNSRLIAGE-PEGKSAIYVLG- 348
Db 378 NMDKLMGNKKIIPFSSI-KALDVAESILSTL--EKSSQN--GYGFGYNN-----IGT 428
Qy 349 IKDRFVINDT-----SLEGNDKILATPAISEAKKEFEKTR 385
Db 429 VKNNSVGEAGEYELIDPVFMKNITSSQAMSLASLAKANAASSDNKKRLSEVATK 485

```

```

RESULT 13
BMPC_BORBU STANDARD; PRT; 353 AA.
AC 050169; Q44859;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Basic membrane protein C precursor.
GN BMPC OR BB0384.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=297;
RA MEDLINE=95080623; PubMed=7968902;
RX Aron L., Aleksun M., Perlée L., Schwartz I., Godfrey H.P.,
RA Cabello F.;
RT "Cloning and DNA sequence analysis of bmpC, a gene encoding a
RT potential membrane lipoprotein of Borrelia burgdorferi.";
RL FEMS Microbiol. Lett. 123:75-82(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=297;
RA MEDLINE=97132632; PubMed=8978084;
RX Aron L., Toth C., Godfrey H.P., Cabello F.C.;
RT "Identification and mapping of a chromosomal gene cluster of Borrelia
RT burgdorferi containing genes expressed in vivo.";
RL FEMS Microbiol. Lett. 145:309-314(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Frazer C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.R., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kariya A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).

```

```

CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U49938; AAC44711.1; -.
DR EMBL: AE001143; AAC66756.1; -.
DR TIGR: BB0384; -.
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp; 1.
DR PROSITE: PS00013; PROKAR LIPOPROTEIN; 1.
KW Membrane; lipoprotein; Signal; Complete proteome.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 353 BASIC MEMBRANE PROTEIN C.
FT LIPID 17 17 N-ACYL DIGLYCERIDE (PROBABLE).
FT VARIANT 179 179 V -> L (IN STRAIN 297).
SQ SEQUENCE 353 AA; 39824 MW; 2E8FF607D6CAB9B0 CRC64;

```

Query Match 6.8%; Score 146; DB 1; Length 353;  
 Best local Similarity 23.7%; Pred. No. 0.035;  
 Matches 84; Conservative 53; Mismatches 131; Indels 86; Gaps 18;

```

Qy 55 VTAGGVYNDNSFNGSGHEAIOQLG-----ALTGEITSVDNSTALEGKY 99
Db 34 VLAHGSFYDKGYNOVDGVKLRDNGIKLITKSLRYPYIEGKRLTYDE--AMEDAY 91
Qy 100 SSLANTKNVNWVSGPHGDAFTRWLKIPEKOLFETKNIILIGI-----DWTPTENVIP 155
Db 92 EVOKNPINLFWLI-GYRFSL-----SVKLSYRPPYIIIGIIDAIFYGDIQ--VPK 139
Qy 156 GRYINLYTEAGWLAGYANASFLAKRPSDPTKRSAYIG--GGISPAVYDFIAGYL 212
Db 140 NS-LAIKFRNEEAFLAGYIAAKMSRKE-----KIGFLTGPMSHVADFKGKF 187
Qy 213 AGIKAMN--LKNSDKRTKITTDKIEINMGFDVODTSTKELEQIA---SKDRPSTLLAV 266
Db 188 AGIFYANPKRLRLVSKKAPSLFD-----KERGKAMALFMYKEDYGVYFPI 232
Qy 267 AGPTEIFSDIIANQND--VLIGVDFDQSLVYTKRKPKFTSLIKNLGYSVFVSLDYLT 325
Db 233 AGITGCGVYAAKELGKRYIVIGLNDQSYT---APQNVTTSTIKDKKIVYISISSEYT- 288
Qy 326 KKSNSRLIAGEFERGKKSATVYLGIKDRFVDA-DTSLGNDKILATPAISEAKK 378
Db 289 ---NNRVFKG-----GIIDRGKIEGVIEIVDPVLNN--RLVDEVIDLEKK 331

```

```

RESULT 14
MERA_STRAU STANDARD; PRT; 547 AA.
AC P08663;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mercuric reductase (Ec 1.16.1.1) (Hg(II) reductase).
GN MERA.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87260937; PubMed=3037534;
RA Laddaga R.A., Chu L., Misra T.K., Silver S.;
RT "Nucleotide sequence and expression of the mercurial-resistance
RT operon from Staphylococcus aureus plasmid p1258.";

```

```
RL Proc. Natl. Acad. Sci. U.S.A. 84:5106-5110(1987).
CC -1- FUNCTION: RESISTANCE TO HG(2+) IN BACTERIA APPEARS TO BE GOVERNED
CC BY A SPECIALIZED SYSTEM WHICH INCLUDES MERCURY REDUCTASE. MERA
CC PROTEIN IS RESPONSIBLE FOR VOLATILIZING MERCURY AS HG(0).
CC -1- CATALYTIC ACTIVITY: Hg + NADP(+) + H(+) -> Hg(2+) + NADPH.
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: HOMODIMER.
CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC OXIDOREDUCTASES CLASS-1.
CC -1- SIMILARITY: CONTAINS 1 HMA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L29436; AAA98245.1; -.
DR PIR; E29504; E29504.
DR HSSP; P11959; 1EBD.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR001934; HMA.
DR InterPro; IPR000815; Hg_reductase.
DR InterPro; IPR001100; pyr_redox.
DR InterPro; IPR004099; pyr_redox_dlm.
DR Pfam; PF00403; HMA; 1.
DR Pfam; PF00070; pyr_redox; 1.
DR Pfam; PF02852; pyr_redox_dlm; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00945; HGRDTRASE.
DR PRINTS; PR00411; PNDRTASE.
DR PROSITE; PS01047; HMA_1; 1.
DR PROSITE; PS01047; HMA_2; 1.
DR PROSITE; PS00846; HMA_2; 1.
DR PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
DR Mercutic resistance; Oxidoreductase; Flavoprotein; FAD; NADP;
KW Mercury; Redox-active center; Metal-binding; Plasmid.
FT DOMAIN 5 68 HMA.
FT NP_BIND 87 117 FAD (ADP PART) (PROBABLE).
FT DISULFID 123 128 REDOX-ACTIVE.
FT NP_BIND 378 388 FAD (FLAVIN PART) (BY SIMILARITY).
FT METAL 544 544 HG(2+) (POTENTIAL).
FT METAL 545 545 HG(2+) (POTENTIAL).
SQ SEQUENCE 547 AA; 58565 MW; BOBEA5FCFA9C049 CRC64;
Query Match 6.3%; Score 136; DB 1; Length 547;
Best Local Similarity 22.6%; Pred. No. 0.26; Mismatches 189; Indels 96; Gaps 19;
Matches 102; Conservative 64;

10 EKSDNQKQTDVSKISGLVNERKSEIMAAKADANKHFGMLMAIVTAGTVNDSF--N 67
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
46 ELSDQIKAKONISAAQYQGEESQPSSENSVDPRGDDVLLITSGCAAFSAITKAN 105
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
68 QSGWEAIDQALATGCG--ETTSVDSSTAELGKYSSILANTKNKYNVLSGFGHDAFTRW 124
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
106 ENGAKVAVERGTVGTCVNIGCVPSKTLRAGEINGLAONNP---FTGLQTSAGADL 161
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
125 LKIPENK-----QLETKENIIILGIDWTDIENVIPYGRYINULTYTEAGWLAGANNSF 179
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
162 AQLTEQKGLVSQMKOE-----YIDL---IEEYGFDLIRGEASF 198
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 LAKRF---PSDPTKRSATVIGGISPAVTFDIAG-----YLAGIKANLKNKSDKTKI- 229
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
199 IDDTIOVNGONITSK--FLATGASPAVPE--IQGMNEVDLITLSALELKEVQRLAVI 256
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
230 -----TTDKIEINLGFVDVDTSTKERLEQJASKDKPSTLLAVAGPLTEIFSIDI----- 278
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
257 GSGYIAELGQMFHNLGTEV---TLMQSRERLFTYDPEISBAIDESLTEGGLNLTGVT 313
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
279 ---ANQNR-----ICVDTSQSLVYTTKTKKFFTSILKNLGYV---FSVLSD 322
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
DB 314 YQKVBONKRSYSITYEVNGQEQVIEADQVLAATGRKPTETLNLMSAGVTKGKGEVLTN 373
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 323 IYTKRSNRLNLAGFEFKKSATVY-----LGIDRFVDIADT--SLGNDKVL 368
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 374 EYLQTSNNRIYAGVYTGPGQFYVAATEGGIVANNALGLAKRKIDLRFPVGVTTNPSTI 433
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 369 ATEAISE--AKPEPEKTKTIPAEVRKTL 396
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 434 ATVGLTEQAKREKGYDVTSVLPDLAVPRAL 464
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
PEDB_STRPY STANDARD; PRT; 498 AA.
ID PEDB_STRPY
AC 099XSL;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable dipeptidase B (EC 3.4.-.-).
DE PEDB OR SPY2066.
GN Streptococcus pyogenes.
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., Moshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -1- CATALYTIC ACTIVITY: Dipeptide + H(2O) -> 2 AMINO ACID.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 034.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006627; AAK34726.1; -.
KW Hydrolyase; dipeptidase; Complete proteome.
SQ SEQUENCE 498 AA; 55499 MW; B9394DC94C19ABAD CRC64;
Query Match 6.0%; Score 129; DB 1; Length 498;
Best Local Similarity 20.5%; Pred. No. 0.61; Mismatches 175; Indels 144; Gaps 23;
Matches 99; Conservative 65;

4 KETTKERSADNQN-----KQTDVSKISGLVNERKSEIMAAKADANKHFG 50
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
63 KDNPAQEKWKDLNSGFEPYLPBHSYRYSAPDVPNKKGVDE-----AGFNE-FGV 112
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 51 NMAIVTAGTYND-----NSFNQSGWEAIDQALATGCG---ITSVDSSTAEL 96
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 113 SMS-ATVASANDALQKIDPYVKNGLAESSMYSVLPSPVAKRGVALIAKIVTEKGA 171
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 97 GYSSILANTKNKYNV---YLSGFQHDADFTRWLAKIPENQQLTEKNIILGIDWTDIENV 153
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 172 CNIVTLAKDK-GIYMEILSGHOYA-----IKRPDKYAAFPNPFYIGHVDFNDKEMTI 225
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 154 PTGRYINLFTYTEAGWLAGYANASFLAKRF--PSDPTKRSATVIGGISPAVTFDIAGY 211
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 226 ASEDVEKAKKASYTEVDKFKH---IAKSYNPPLDANRSRFSFG----- 268
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 LAGIKANLKNKSDKTKIT-----TDKIEINLGFVDVDTSTKER----- 250
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
```



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:31:16 ; Search time 50.25 Seconds  
(without alignments)  
1456.256 Million cell updates/sec

Title: US-09-676-249A-4

Perfect score: 2155  
Sequence: 1 MMDKETEKEKSADNOKQI.....KQPDKQKSLDKLITDINNL 423

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SP\_ARCHAEA:\*
- 2: SP\_BACTERIA:\*
- 3: SP\_FUNGI:\*
- 4: SP\_HUMAN:\*
- 5: SP\_INVERTEBRATE:\*
- 6: SP\_MAMMAL:\*
- 7: SP\_MHC:\*
- 8: SP\_ORGANELLE:\*
- 9: SP\_PHAGE:\*
- 10: SP\_PLANT:\*
- 11: SP\_RODENT:\*
- 12: SP\_VIRUS:\*
- 13: SP\_VERTEBRATE:\*
- 14: SP\_UNCLASSIFIED:\*
- 15: SP\_VIRUS:\*
- 16: SP\_BACTERIAP:\*
- 17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 693.5 | 32.2        | 457    | 2     | Q48902 mycoplasma  |
| 2          | 552   | 25.6        | 461    | 16    | Q980L5 mycoplasma  |
| 3          | 391   | 18.1        | 465    | 2     | Q9X775 mycoplasma  |
| 4          | 384.5 | 17.8        | 428    | 2     | Q52311 mycoplasma  |
| 5          | 382.5 | 17.7        | 428    | 2     | Q9RGX5 mycoplasma  |
| 6          | 379.5 | 17.6        | 428    | 2     | Q32417 mycoplasma  |
| 7          | 379   | 17.6        | 429    | 2     | Q9RGX6 mycoplasma  |
| 8          | 379   | 17.6        | 429    | 2     | Q9RGX4 mycoplasma  |
| 9          | 377.5 | 17.5        | 428    | 2     | Q9RGX7 mycoplasma  |
| 10         | 374.5 | 17.4        | 428    | 2     | Q9RGX3 mycoplasma  |
| 11         | 373.5 | 17.3        | 428    | 2     | Q9RGX6 mycoplasma  |
| 12         | 228   | 10.6        | 349    | 2     | Q9EV91 clostridium |
| 13         | 220.5 | 10.2        | 350    | 16    | Q9GZM4 streptococc |
| 14         | 203.5 | 9.4         | 350    | 16    | Q9GPM9 lactococcus |
| 15         | 202   | 9.4         | 350    | 16    | Q97RHO streptococc |
| 16         | 192   | 8.9         | 357    | 16    | Q97L60 clostridium |

|    |       |     |      |    |                    |
|----|-------|-----|------|----|--------------------|
| 17 | 182   | 8.4 | 348  | 2  | Q9AK41 streptomyc  |
| 18 | 178.5 | 8.3 | 359  | 16 | Q9XWV7 thermotoga  |
| 19 | 174   | 8.1 | 342  | 16 | Q9A1P7 streptococc |
| 20 | 173   | 8.0 | 357  | 16 | Q92BW7 listeria in |
| 21 | 170   | 7.9 | 264  | 2  | Q31358 borrelia ga |
| 22 | 168.5 | 7.8 | 516  | 16 | Q9PRD3 ureaplasma  |
| 23 | 168   | 7.8 | 341  | 2  | Q9FED3 borrelia bu |
| 24 | 164   | 7.6 | 347  | 2  | Q9AK42 streptomyc  |
| 25 | 162.5 | 7.5 | 525  | 16 | Q9PQR5 ureaplasma  |
| 26 | 161   | 7.5 | 339  | 2  | Q9S6C1 borrelia bu |
| 27 | 161   | 7.5 | 339  | 2  | Q87960 borrelia bu |
| 28 | 160   | 7.4 | 325  | 2  | Q9R776 borrelia bu |
| 29 | 157   | 7.3 | 339  | 2  | Q9FED2 borrelia bu |
| 30 | 156   | 7.2 | 325  | 2  | Q9R777 borrelia bu |
| 31 | 150   | 7.0 | 524  | 16 | Q9PRC9 ureaplasma  |
| 32 | 148   | 6.9 | 329  | 16 | Q9RGW4 rhizobium l |
| 33 | 145   | 6.7 | 264  | 2  | Q31359 borrelia ga |
| 34 | 143.5 | 6.7 | 243  | 2  | Q31361 borrelia ga |
| 35 | 143   | 6.6 | 353  | 2  | Q93V09 borrelia bu |
| 36 | 140   | 6.5 | 379  | 16 | Q9RSQ4 deinococcus |
| 37 | 137.5 | 6.4 | 889  | 2  | Q9L961 moraxella c |
| 38 | 134   | 6.2 | 539  | 16 | Q9PQ10 ureaplasma  |
| 39 | 133.5 | 6.2 | 1786 | 5  | Q9UDP0 plasmidum   |
| 40 | 132.5 | 6.1 | 330  | 16 | Q926D6 rhizobium m |
| 41 | 132   | 6.1 | 326  | 16 | Q92X35 rhizobium m |
| 42 | 131.5 | 6.1 | 1558 | 5  | Q96275 plasmidum   |
| 43 | 127   | 5.9 | 912  | 2  | Q9XD52 moraxella c |
| 44 | 127   | 5.9 | 2045 | 16 | Q9AOK5 streptococc |
| 45 | 126.5 | 5.9 | 894  | 2  | Q9L962 moraxella c |

ALIGNMENTS

RESULT 1  
ID Q48902 PRELIMINARY; PRT; 457 AA.  
AC Q48902;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE AG 243-5 PROTEIN PRECURSOR.  
OS Mycoplasma arginini.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2094;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96163149; PubMed=8551970;  
RA Ushio S., Iwaki K., Tanai M., Ohta T., Fukuda S., Sugimura K.,  
RA Kurimoto M.;  
RT "Metastatic promoting activity of a novel molecule, Ag 243-5 derived  
RT from Mycoplasma and the determination of the complete nucleotide  
RT sequence.";  
RL Microbiol. Immunol. 39:393-400(1995).  
DR EMBL: D16674; BAA04082.1; -  
DR InterPro: IPR003760; Bmp.  
DR Pfam: PF02608; Bmp; 1.  
KW Signal.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 457 AG 243-5 PROTEIN.  
SQ SEQUENCE 457 AA; 50789 MW; 9AEEBA11620CB22F CRC64;

Query Match 32.2%; Score 693.5; DB 2; Length 457;  
Best Local Similarity 39.0%; Pred. No. 1,9e-32;  
Matches 188; Conservative 81; Mismatches 141; Indels 41; Gaps 14;  
QY 5 ETTRKEKSADNOKQITDVSKISGLVNERSEIMAAKADNKHGMLMATTAGGTYNDN 64  
DB 28 ETTRKEGKIT-----RIFD-----NSFVDRQAELEKA-----KNDFVTYVLTITGGTVQDK 73  
QY 65 SFNOSGWEAI---QDGLALTG--GEITSVDSSTAELGKYSILANTNKNWVVLGFOHGD 119

```

Db 74 SNNOSIWEVLEHYDOIERTKTMLDRVSOETNNQSELIGKFKFLNGKNKVMWILTGQOQO 133
QY 120 AETRMKLIPEN-----KOLFEKNIIILIGIDW---TDENVIPTGRYINLTYKTEBAGWL 171
Db 134 EEPKRLKOTDSNGKKYSODLAEKKYIYAVBDLSKEDDLKAGHIFSLYKTEBAGVI 193
QY 172 AGYANNSFLAKKFPSPDPTKRSAYIVGGISPAVDTFIAGYLAGIKAMNLSKDKTKIT 231
Db 194 AGYASSKFLAYKFPNDEAKRTIAPFGCGHAGVTDFIAGFLAGIKAYNNNDNPJAKVTISD 253
QY 222 DKIEINLGFVDOTSTKERLEQIASKDKRSTLLAVAGPLTEIFSDI--ANONDRLLIYV 289
Db 254 NNINIDTGF-ISNDKTAFINCIYVKS--SLVLPVAGSLTSSVVAIKKSNKDTKRLIGV 310
QY 290 DDDQSLVYTKTKNKFPTSLIKNLGYSVFSVLSLDLYTKKSNHNLAGEFGK---KSATV 345
Db 311 DTDQSKITSPA-TYFTSTLEKHLGRTIYVLDIMLKEDSKFSLGFRFKLTNPANATV 369
QY 346 YGIGKDFVADIADTSLSEGNDRKLATEAISANKFEETKRTIPAEVRKTEIPEM--P 402
Db 370 YGIGSDDFVCGNSSTVADADKVKAGQEFLEATADFKQIQOANPT-NYKSVLGIPFWLIND 428
QY 403 DKQPKQOESL 413
Db 429 NDAKDNKASL 439

RESULT 2
QY 0980L5 PRELIMINARY: PRT: 461 AA.
AC 0980L5:
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE ABC TRANSPORTER XYLOSE-BINDING LIPOPROTEIN.
GN MYPU_3460.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molllicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OAB CTIP;
RX MEHLIN=21267165; PubMed=11353084;
R Moszer I., Dybvig K., Wroblewski H., Viart A., Rocha E.P.C.,
Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
Mycoplasma pulmonis."
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AL445564; CAC13519.1; -.
DR MyPulast; MYPU_3460; -.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
KW Complete proteome.
SQ SEQUENCE 461 AA; 51096 MW; 6724D4D820809CE4 CRC64;

Query Match 25.6%; Score 552; DB 16; Length 461;
Best Local Similarity 31.2%; Pred. No. 2.9e-24;
Matches 133; Conservative 82; Mismatches 161; Indels 50; Gaps 11;
QY 13 ADNOKQIT--DVSKISGLVNER-----KSEIMAAKADANKHFGIMNAVYAGGV 61
Db 28 AONPKKTSNDSKKITDLSQKVEYTEROKYENKIKQASLETQK-----VVLITADGNI 82
QY 62 NDNSFGWEAIIQQLGALTGGEITS-----VDSSTAELEGKYSSILANTNK 107
Db 83 DDKSFNQOVYESOKTLKDFVDKAYKSONKEANQHKLDNYINSAYVDEQNKVVALDGRY 142
QY 108 NWVYLSGFOHDAFTRMKLIPENKOLFTEKNIIILIGIDMTDEN-VIPGRKINLTYTE 166

```

```

Db 143 TTVILTGFOQGEIENFLNDENNLRFEKNKVIIGVDPANPANSKIFQSLISLTFTE 202
QY 167 EAGWLAGYANASFLAKKFPSPDPTKRSAYIVGGISPAVDTFIAGYLAGIKAWN--LKNSD 224
Db 203 EAGWQAGVYASAPFLGTKYANNNAKRAISAFGGDAGVYDFLNGFPEGIRAMNSAEKAN 262
QY 225 KTKKITTDKIEINLGFVDOTSTKERLEQIASKDKRSTLLAVAGPLTEIFSDI--ANQ 281
Db 263 KKKIVISENLVYDTGF-IPNAEKNEVSNVYETGKSTISLPVAGDPFGVYVDLKRQSD 321
QY 282 NDRYLIVDTQDSLTYTKNKFPTSLIKNLGYSVFSVLSLDLYT-----KKSNSNL 333
Db 322 EDRFTYGVDTQDSLSTFNDSKRFSTSYKNIAPVYQILALLTLDESDYILKEGNDKFL 381
QY 334 AGFEFGKSAVYVLCIKDFVADIADTSLSEGNDRKLATEAISEA--KKEFEETKTIPEE 391
Db 382 GS---NPKNLVYKRLSKAFVNIITSRKVESIKTQADISIQAIKMANPANSKIEREM 438
QY 392 VKRTLE 397
Db 439 TNGDLE 444

RESULT 3
QY 09X775 PRELIMINARY: PRT: 465 AA.
AC 09X775:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE P48 MEMBRANE LIPOPROTEIN PRECURSOR.
GN P48.
OS Mycoplasma agalactiae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molllicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2110;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M7;
RX MEDLINE=20002620; PubMed=10531294;
RA Rosati S., Pozzi S., Robino P., Montinaro B., Conti A., Fadda M.,
Pittau M.;
RT "P48 major surface antigen of Mycoplasma agalactiae is homolog to a
malp product of Mycoplasma fermentans and belongs to a selected family
of bacterial lipoproteins."
RL Infect. Immun. 67:6213-6216(1999).
DR EMBL; AJ132423; CAB43718.1; -.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
KW Signal; Lipoprotein.
FT SIGNAL 1 22
FT CHAIN 23 465 P48 MEMBRANE LIPOPROTEIN.
SQ SEQUENCE 465 AA; 51149 MW; 60AD5448CFEB03C96 CRC64;

Query Match 18.1%; Score 391; DB 2; Length 465;
Best Local Similarity 31.2%; Pred. No. 5.9e-15;
Matches 119; Conservative 76; Mismatches 128; Indels 58; Gaps 18;
QY 55 VTAGTVADNSFNQSGWEAIIQQLGALTGGEITSVDSSTAELEGK-----YSSL 102
Db 70 ITDEGSVHDESNQSGWEAIVHVSVEIGLDKRAQV-SGNKNLRNKYEPKQQLLEAVYKNA 128
QY 103 ANTNNWVVLGFOGDAFTRMKLIPENK-OLFEKNIIILIGIDWT-----DENVY----- 152
Db 129 IDSGRYIVLCGFTQASL---VGLDENYIKKIDNNIIFITVDFNLTEDDANKYTRIK 185
QY 153 -IPTGRYINLTYYKTEAGWLAGYANASFLAKKFPSPDPTKRSAYIVGGISPAVDTFIAGY 211
Db 186 KIGEGHLVAVIPFDTRQAAVIAGBALADYFSQYKQPKRTIGAGGIPWPAVSDFIAGT 245
QY 212 LAGIYAMWLKNSDKTKITTTDKIEINLGFVDOTSTKERLEQIASKDKRSTLLAVAGPL 270

```

|    |     |   |                              |     |
|----|-----|---|------------------------------|-----|
| Db | 246 | FOGIIIDMKNREPEAKTSLSNETIELNTLF---                         | TSGTPOATTAINSVYKAAASYPVAGSLT | 302 |
| Qy | 271 | TEITSDD--IANQDRILIGYDIDQSLVYKIKR----                      | KFFSILANLGSYFVSVALD          | 323 |
| Db | 303 | TDTKEIKRLKLD--DKELIGDADQ-----                             | KNALGHRIFESYMKLIGQAAVYNIILAD | 354 |
| Qy | 324 | YTKRSNSRNL-AGFEPEKRSAT--VYLIGKD---                        | RFDIADITS--LEBNQDKLAKTEALSEA | 376 |
| Db | 355 | YSKSEONIDLOPGEIEKKNKSTPVFVGEGYDTEKQYGVATSGLLDKNDEIANKLKDA | 414                          |     |
| Qy | 377 | KKEEETKTTPAEVEKRTLE                                       | 397                          |     |
| Db | 415 | TATYVQK-KT-----EIQKSLK                                    | 430                          |     |

|    |  |                 |                         |
|----|--|-----------------|-------------------------|
|    | RESULT   | 4               |                         |
| KM | 052311   |                 |                         |
| AC | 052311   | PREDIMINARY;    | PRT; 428 AA.            |
| DT | 01-JUN-1998  | (TREMBLrel. 6,  | Created)                |
| DR | 01-JUN-1998  | (TREMBLrel. 06, | Last sequence update)   |
| PT | 01-JUN-2001  | (TREMBLrel. 17, | Last annotation update) |
|    | MEMBRANE LIPOPROTEIN P48V.                                   |                 |                         |
| OC | Mycoplasma fermentans.                                       |                 |                         |
| CC | Bacteria; Firmicutes; Bacilli/Clostridium group; Mollicutes; |                 |                         |
| OX | NCBI_Taxid=2115;   |                 |                         |
| RN | [1]  |                 |                         |
| RP | SEQUENCE FROM N.A.   |                 |                         |
| RC | STRAIN=PG18;   |                 |                         |
| RA | Ravadi G., Dyer K., Dujeancourt A.;                          |                 |                         |
| DR | Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.      |                 |                         |
| DR | EMBL: AF036106; AAB99740.1; -                                |                 |                         |
| DR | InterPro: IPR003760; Bmp.                                    |                 |                         |
| DR | Pfam: PF02608; Bmp; 1.                                       |                 |                         |
| KM | Lipoprotein.   |                 |                         |
| CM | SEQUENCE 428 AA; 47862 MW; D9506E817E330EDA CRC64;           |                 |                         |

[illegible]

| RESULT | 5  |              |           |  |
|--------|--|--------------|-----------|--|
| 09RGX5 |  |              |           |  |
| ID     | 09RGX5   | PRELIMINARY; | PRT;      | 428 AA.                                |
| DT     | 01-MAY-2000 (TREMBLrel. 13, Created)                                   |              |           |  |
| DT     | 01-MAY-2000 (TREMBLrel. 13, Last sequence update)                      |              |           |  |
| DT     | 01-JUN-2001 (TREMBLrel. 17, Last annotation update)                    |              |           |  |
| DE     | MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.                       |              |           |  |
| GN     | MALP.  |              |           |  |
| OC     | Mycoplasma fermentans.   |              |           |  |
| OC     | Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;          |              |           |  |
| OC     | Mycoplasmataceae; Mycoplasma.  |              |           |  |
| OX     | NCBI_TaxID=2115;   |              |           |  |
| RN     | [1]  |              |           |  |
| RP     | SEQUENCE FROM N.A.   |              |           |  |
| RC     | STRAIN=SK5;  |              |           |  |
| RX     | MEDLINE=99115554; PubMed=9916088;                                      |              |           |  |
| RA     | Calcutt M.J., Kim M.F., Kaipras A.B., Muhlradt P.F., Wise K.S.;        |              |           |  |
| RT     | "Differential posttranslational processing confers intraspecies        |              |           |  |
| RT     | lipoprotein of a major surface lipoprotein and a macrophage-activating |              |           |  |
| RL     | infect. Immun. 67:760-771(1999).                                       |              |           |  |
| DR     | EMBL; AF099211; AAD16395.1; --   |              |           |  |
| DR     | InterPro; IPR003760; Bmp.  |              |           |  |
| DR     | Pfam; PF02608; Bmp; 1.   |              |           |  |
| KW     | Signal; Lipoprotein.   |              |           |  |
| FT     | SIGNAL   | 1            | 24        | POTENTIAL.                             |
| FT     | CHAIN  | 25           | 428       | MACROPHAGE ACTIVATING LIPOPROTEIN-404. |
| SO     | SEQUENCE   | 428 AA;      | 47835 MW; | D03F047A2B1460 CRC64;                  |

|        |   |                  |                                 |            |             |  |
|--------|---|------------------|---------------------------------|------------|-------------|--|
|        | Query Match   | 17.7%;           | Score 382.5;                    | DB 2;      | Length 428; |  |
|        | Best Local Similarity   | 27.0%;           | Pred. No. 1,6e-14;              |            |             |  |
|        | Matches 116;  | Conservative 87; | Mismatches 156;                 | Indels 71; | Gaps 17;    |  |
| OY     | 31 NERKSEIMAAKADANKH-----FGINMAIVTAGTVNDNSFNQSGWEAI                   | 74               |                                 |            |             |  |
|        | :   :   :   :   |                  | :   :   :   :   :               |            |             |  |
| Db     | 27 NNDESNISFEKEDISKYTITNANGKVYKNAELLKLPLVLTIDEGKIDDKSFNQSAFEL         | 86               |                                 |            |             |  |
| OY     | 75 OOLGALGTGELTSVVSSPTELDEGKYSSLANTKKNNWVLSCFOHGDAFTWLKIPENKOLF       | 134              |                                 |            |             |  |
|        | : :   :   :   :   :   :   :   |                  | : :   :   :   :   :   :   :     |            |             |  |
| Db     | 87 KAINRQTGEINSVEPS--SNFESAYNSALSAGHKIWLVLGFKHQOOSIKQY--IDAAREL       | 143              |                                 |            |             |  |
| OY     | 135 TEKNTIIIGIDMTDENVIPTGRYINLVTKTEBGAGLVANASFLAKKEPSPDTRRSAI         | 194              |                                 |            |             |  |
|        | :   :   :   :   :   :   :   :   :                                     |                  | :   :   :   :   :   :   :   :   |            |             |  |
| Db     | 144 ERNOIKITIIDE-DIETEKX--FYSLQPNIKESAFTGYALASWLSIQ--DESRUYVA         | 197              |                                 |            |             |  |
| OY     | 195 YVGIGISPAPVTDFINGVLAGIRAMWLKNSDKRTKI--TUDKIEINLGFDVDQSTKERLEQ     | 253              |                                 |            |             |  |
|        | :   :   :   :   :   :   :   |                  | :   :   :   :   :   :   :   :   |            |             |  |
| Db     | 198 SPGGGAFPGVTFNFNGFPAGKILLYONKH--KSKIKHTSPVLDLSF-----TAGCKMNT       | 250              |                                 |            |             |  |
| OY     | 254 IAS-----KRPSTLLAVAGBLEIFSDIIIANONDRLYGVDTPQSLVYYTKTKA             | 303              |                                 |            |             |  |
|        | : :   :   :   :   :   :   :   :   :   :                               |                  | : :   :   :   :   :   :   :   : |            |             |  |
| Db     | 251 VYNNVLSSTPADVKKNPHILSVAGBAT--FETVRLANGGVYIVIGDSQOGMI--ODKR        | 306              |                                 |            |             |  |
| OY     | 304 FFTSLTKNLNGVSVEVSLDL-----YRKSSRMLAPEPECKKATAYYLGIKIDKF            | 353              |                                 |            |             |  |
|        | :   :   :   :   :   :   :   |                  | :   :   :   :   :   :   :   :   |            |             |  |
| Db     | 307 ILTSVQLKHQRQVAVETLDDLLEKEBGPYPYVVKDKAKAKMSHFSTQ-----KEEW          | 358              |                                 |            |             |  |
| OY     | 354 VDIAOTSLLEGNDKUL-AATEAISAEAKKEFEERTKTTPAEAEVRKTLLEIREMPD-KOPDKOOE | 411              |                                 |            |             |  |
|        | : :   :   :   :   :   :   :   :   :   :                               |                  | : :   :   :   :   :   :   :   : |            |             |  |
| Db     | 359 IGVAENHHSNNEEQAKINKIKEALKMFKE----LPEDVKYKINDSKALDKGNKDINVSE       | 414              |                                 |            |             |  |
| OY     | 412 SLDKLTIDTN 421  |                  |                                 |            |             |  |
|        | :   :   :   :   |                  |                                 |            |             |  |
| Db     | 415 RLEAIYSAIN 424  |                  |                                 |            |             |  |
| RESULT | 6   |                  |                                 |            |             |  |
| 032417 |   |                  |                                 |            |             |  |

[illegible]

| QY                        | 354  | VIADPSTLEGNCKL-ATBAISEAKKEPEKRTKTIIPAEVRKLTLEIPMPD-KQDPKOE   | 411                                    |
|---------------------------|--|--|--|
| DB                        | 339  | IYVAENHNSNTEQAKINKKIKKAIKMKFE---LPEDFVKTIINSKDKALDGNKIDNWSE  | 414                                    |
| QY                        | 412  | SLDKLTITDIN 421  |  |
| DB                        | 415  | RLAIIISALN 424   |  |
| RESULT                    | 7  |  |  |
| Q9RGX6                    | PRELIMINARY;   | PRT;   | 429 AA.                                |
| AC                        | Q9RGX6   |  |  |
| DT                        | 01-MAY-2000 (Tremblrel. 13, Created)                                 |  |  |
| DT                        | 01-MAY-2000 (Tremblrel. 13, Last sequence update)                    |  |  |
| DT                        | 01-DEC-2001 (Tremblrel. 19, Last annotation update)                  |  |  |
| DE                        | MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.                     |  |  |
| GN                        | MALP.  |  |  |
| OS                        | Mycoplasma Incoognitus.  |  |  |
| OC                        | Bacteria; Firmicutes; Bacillus/Clostridium group; Molllicutes;       |  |  |
| OC                        | Mycoplasmataceae; Mycoplasma.  |  |  |
| OK                        | NCBI_TaxID-2101;   |  |  |
| RN                        | [1]  |  |  |
| RP                        | SEQUENCE FROM N.A.   |  |  |
| RC                        | STRAIN-INCOGNITUS;   |  |  |
| RX                        | MEDLINE-99115554; PubMed-9916088;                                    |  |  |
| RA                        | Calcutt M.J., Kim M.F., Kaipas A.B., Muhlradt P.F., Wise K.S.;       |  |  |
| RT                        | "Differential posttranslational processing confers intraspecies      |  |  |
| RT                        | variation of a major surface lipoprotein and a macrophage-activating |  |  |
| RT                        | lipopeptide of Mycoplasma fermentans.";                              |  |  |
| RL                        | Infect. Immun. 67:760-771(1999).                                     |  |  |
| DR                        | EMBL: AF099210; AAD16394.1; -  |  |  |
| DR                        | InterPro: IPR003760; Bmp.  |  |  |
| DR                        | Pfam: PF02608; Bmp; 1.   |  |  |
| DR                        | Signal; Lipoprotein.   |  |  |
| KW                        | SIGNAL   | 1 24   | POTENTIAL.                             |
| FT                        | CHAIN  | 25 429   | MACROPHAGE ACTIVATING LIPOPROTEIN-404. |
| FT                        | SEQUENCE   | 429 AA; 47961 MW; B6D09A8812AC3171 CRC64;                    |  |
| Query Match               | 17.6%;   | Score 379;   | DB 2; Length 429;                      |
| Best Local Similarity     | 27.0%;   | Pred. No. 2.6e-14;   |  |
| Matches 117; Conservative | 88;  | Mismatches 151;  | Indels 78; Gaps 18;                    |
| QY                        | 31   | NERKEIIMAAKADAKH-----FGLNMAIYTAGTVNDNSNOSGWEAI               | 74                                     |
| DB                        | 27   | NNDESNISFKEKDISKYTTTANGQOVVKNMAELLKLPVLTIDECKIDKSNOSAFAL     | 86                                     |
| QY                        | 75   | QOILGTLGELTISVDSSTAELEGGKYSILANTKNVWVLSGFOGDAFTRWLIKIPENKOLF | 134                                    |
| DB                        | 87   | KAINKOTGELINNVPS-SNFESATNSALSGHKITWLVNGFHHOOSIOY--IDAHREL    | 143                                    |
| QY                        | 135  | TEKNIIIGIDWTDENVYIPGRYINLTVKTEAGWLAGVANASFLARKFPSPTRKSAI     | 194                                    |
| DB                        | 144  | ERNQIKIIGIDF-DIETEXKW--FYSLOFNIKESAFITGYALASMSQ--DESKRYVA    | 197                                    |
| QY                        | 195  | VIGGISPAWDFINGYLAGIKANMLKSDKTKI-TIDKIEINLGFVQDPTSTERLEQ      | 253                                    |
| DB                        | 198  | SFGGAGFAGVYTFENGGFKGILYYQKH--NSSKIYHTSPVXLDSGF-----TAGEKNMT  | 250                                    |
| QY                        | 254  | IAS-----KRPSTLLAVAGPLTEIFSDI-IANONDIYLLGVDPDQSLVYTKTKN       | 302                                    |
| DB                        | 231  | VINNVLSSTPADYKKNPHVILSYAGPAI--FETVYLANKGQYVIGVSDQGMT--QDKD   | 306                                    |
| QY                        | 303  | KFTSILANKLGSYFVSVDL-----YTKSNSRNLAGFEFGKKSATVYLGIDR          | 352                                    |
| DB                        | 307  | RLTFSVLKHIQAVYETLLDILKEEGKPPVVDKADKKWMSFGQ-----KEK           | 358                                    |
| QY                        | 353  | FVDIADTSLGDKKL-ATEAISEKKFE---EKTITIPAEVYRKLTLEIPMDKQD        | 407                                    |
| DB                        | 359  | WIGVAENFNSNTEQAKINKKIKKAIKMKFELPDEPFVKYIINSKDKVLE-----DGNKID | 411                                    |







QY 20 ITDVSISGLVNERKSEIMAAADANKHGLNMAITVAGTVNDSPNOSGWEALIOQLGA 79  
 DB 13 LASVAVLACGRSHDAAGSKAKTD-----LKAAIYTEIGVVDREGFNOSAMEGLQSGWK 66  
 QY 80 LG-----GEITSVDSSTHELEGGKYSGLANTKNWVLSGFQHGAFTRMLKIPENKOLF 134  
 DB 67 ENNLKKTGYTYFQSSASADYTTNNASDAQGKTLFGFSLQDTSAAK---NN--- 120  
 QY 135 TEKNITILIDWTDENVIPTGRYINLTYKTEEAGMLAGYANASFLAKFPSPDPTKRSAT 194  
 DB 121 PKNFPIYDSYVKDKNV-----ASATFADNESATLAGVAAK-----ATTKNTI 165  
 QY 195 -VIGGICSPAVTDFIAGYLAGIKAMNLSKDKTKITTDKIEINLGFVDPTSTKERLEQ 253  
 DB 166 GFIGGMSQDVITREFEKYBAGAKSVN-----PDIKVDYVAGSFSDAKGTIAA 215  
 QY 254 IASKDKPSLLAVAGPL-TEIFSDIA-----NQNDR-YLIGVDIDQSLV--YT-----KT 300  
 DB 216 AMYGAGDDVYQACGCGVGFSEAKALNSTKNADKVVIVIGDDQEKLGKYSKDGAD 275  
 QY 301 KKKFPTSLKNLGYSVFVLSLDLYTKKNSRNLAG--FEFGKKSATVYLGIRDFVDIAD 358  
 DB 276 SNFVLVSTIKEG-----NVKDIADKTKDGKFPGGITVYDILKNGVNLGL-----DSAN 326  
 QY 359 TSLEGNDKRLATEAISEAKKEFEKTKTIPAE 390  
 DB 327 SEIK-----DAVAKAKADIIDGKITVPSK 350

RESULT 15

Q97RHO PRELIMINARY; PRT; 350 AA.  
 AC Q97RHO;  
 DT 01-OCT-2001 (TREMBlrel. 18, Created)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE LIPOPROTEIN.  
 GN SP0845.  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TIGR4;  
 RC MEDLINE=21357209; PubMed=11463916;  
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
 Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,  
 Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,  
 Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
 Holtzapple E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,  
 McDonald L.A., Feldlyum T.V., Angiolli S., Dickinson T., Hickey E.K.,  
 Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
 Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.,  
 "Complete genome sequence of a virulent isolate of Streptococcus  
 pneumoniae."  
 RT Science 293:498-506(2001).  
 RL EMBL: AE007390; AAK74976.1. -  
 DR EMBL: AE007390; AAK74976.1. -  
 DR TIGR: SP0845. -  
 DR InterPro: IPR001114; Adenylsucc\_synth.  
 DR InterPro: IPR003760; Bmp.  
 DR InterPro: IPR000217; Tubulin.  
 DR Pfam: PF02608; Bmp. 1.  
 DR PROSITE: PS00227; TUBULIN; UNKNOWN\_1.  
 KW Complete proteome.  
 SO SEQUENCE 350 AA; 36746 MW; CBD767B7EF501E69 CRC64;

Query Match 9.4%; Score 202; DB 16; Length 350;  
 Best Local Similarity 25.7%; Pred.No.0.00033;  
 Matches 98; Conservative 55; Mismatches 145; Indels 84; Gaps 19;

QY 33 RKSEIMAAKADANKHGLNMAITVAGTVNDSPNOSGWEALIOQLGA---LNGEITSVD 89

DB 25 RSSRNAASSSDVK-----TKAAIYTDGVDKSPNOSAMEGLQAMGKEHNLKXNGFTYF 80  
 QY 90 SSTAEE-----GKSSSLANTKNWVLSGFQHGAFTRMLKIPENKOLFTEKNTI 140  
 DB 81 OSTSEADYANNLQQAAGSYNLIFFV-----GFLANNA-----VKDAAEHTDLNV 126  
 QY 141 ILGIDWTDENVIPTGRYINLTYKTEEAGMLAGYANASFLAKFPSPDPTKRSATVYIGGI 200  
 DB 127 LIDDVYKDKNV-----ASVTFADNESGYLAGVAAK-----TTKTQVGVGCI 171  
 QY 201 -SPAVTDFIAGYLAGIKAMNLSKDKTKITTDKIEINLGFVDPTSTKERLEQIASDK 259  
 DB 172 ESEVISRFEAGFKAGV-----ASVDPISIKVQVDYAG-SFG-DAAGKTIAAAQVAAAGAD- 223  
 QY 260 PSTLLAVAGPL-TEIFSDIANQNDR-----YLIGVDIDQSLV--YT-----KT 306  
 DB 224 -IVYQVAGGTGAGVFAEAKSLNESRPNENKVVVIGVDDQAEKRYTSKDGESNFVLY 281  
 QY 307 SILKNLGYSVFVLSLDLYTKKNSRNLAGEFEFGKKSATVYLGIRDFVDIADTSLEGNDK 366  
 DB 282 STLKOVGTIVKDI-----SNKAERGFPGGQ---VIVYSLKDKGYDLAVTNLSEEG 330  
 QY 367 KLAATEAISEAKKEFEKTKTIP 368  
 DB 331 K-----AVEDAKAKILDGSVKVP 348

Search completed: July 18, 2002, 15:31:18  
 Job time: 125 sec





**THIS PAGE BLANK (USPTO)**



```

Db      3 RSSRMASSDVK-----TKAIVTDTGVDKSFNOSAMEGIQANGKEHNLSKDNCFITYE 58
QY      90 SSTABLE-----GKYSILANTNKNWVLSGFQHDGDAFTRLKIPENKOLFTEKNII 140
Db      59 OSTSEADYANNLQQAAGSYNLIIFGV-----GFLANNA-----VDAAKEHIDLNVY 104
QY      141 ILGIDMTENVIPGRININTLYKTEEGWLAGVANASFLAKKPPSDPKSAIYIGGI 200
Db      105 LIDVYIKDQKNV-----ASYTFADNESGYLAGVAAAK-----TTKTQVGFVGGI 149
QY      201 -SPAVTDFIAGLAKIKAMNLSKDKTKITTDKIEINLGFVDVDTSTKERLEQJASKDK 259
Db      150 ESEVTSREDFAGKACV-----ASVDPSTIKVVDYAG-SFG-DAAGKTIAAQVYAGAD- 201
QY      260 PSTLAVAGPL-TEIFSDIINONDR-----YLIGVDTQSL-VYT-----TKRKFFT 306
Db      202 -IYQVAGGAGGAFEAKEKSLNESRPNENKRWYIGVDRDQAEKYSKQKESNFVLY 259
Db      307 SILKNLGYSVFSVLSDLTYKKSNSRNLAGFERGKKSATVYLGIDRFVADIADTSLEGNDK 366
Db      260 SLKQVGTIVDI-----SNKAERGFPGGQ---VIVYSLKDKGYDLAVTNLSEEGK 308
QY      367 KLTAEIASEAKKEFEKTKTIP 388
Db      309 K----AVEDAKAKILDGSVKVP 326

```

```

RESULT 2
US-08-396-957A-5
: Sequence 5, Application US/08396957A
: Patent No. 5780041
: GENERAL INFORMATION:
: APPLICANT: SIMPSON, WARREN; SCHWAN, TOM G.
: TITLE OF INVENTION: ANTIGENIC PROTEINS AND
: TITLE OF INVENTION: GENES ENCODING SAME OF BORRELIA BURGDORFERI.
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
: STREET: 345 PARK AVENUE
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: USA
: ZIP: 10154
: COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY DISK
: COMPUTER: IBM PC COMPATIBLE
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORDPERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/396,957A
: FILING DATE: 01-MAR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/020,245
: FILING DATE: 19-FEB-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/664,731
: FILING DATE: 05-MAY-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/487,716
: FILING DATE: 05-MAY-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: DOROTHY R. AUTH
: REGISTRATION NUMBER: 36,434
: REFERENCE/DOCKET NUMBER: 2026-4018054
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 758-4800
: TELEFAX: (212) 751-6849
: TELEEX: 421792
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 341

```

```

: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: HYPOTHETICAL: No
: ORIGINAL SOURCE:
: ORGANISM: Borrelia burgdorferi
: STRAIN: Sh-2-82
: INDIVIDUAL ISOLATE:
: DEVELOPMENTAL STAGE:
: HAPLOTYPE:
: TISSUE TYPE:
: CELL TYPE:
: CELL LINE:
: ORGANELLE:
: FEATURE:
: NAME/KEY: p39a
: LOCATION:
: IDENTIFICATION METHOD:
: OTHER INFORMATION: p39a protein sequence
US-08-396-957A-5

```

```

Query Match      8.3%; Score 178; DB 1; Length 341;
Best Local Similarity 26.9%; Pred. No. 1.1e-08;
Matches 101; Conservative 58; Mismatches 124; Indels 92; Gaps 25;

```

```

QY      44 ANKHFGLNMAIYTAGTYVAGTVDNSFNOSGWEAIQOGLALTGGETTSYDSTAIEGKYS- 101
Db      24 SSKKIKISMLV---DGVLDKDSFNSSANEALRLKDPENIEEVS--CAISGYSSYV 78
QY      102 --LANTNKN---VWVLSFGHGAFTFRLKIPENKOLFTEKNII--ILGIDMTENVY 153
Db      79 SLDLMLKRNGLSGLV-LVGYMLTDA--SLVSESEPKI--SGYIDPIYGDVQVPEKLI 133
QY      154 PGRYINLTYTEEGWLAGY--ANASFLAKKFPSPDKRBAIYG--GGISPAVTD-FI 208
Db      134 A-----VVFVEQCAFAGYIAAKKSPSGK-----IGFGMKMGKNTVDADR 174
QY      209 AGYLAGIKAMNLSKDKTKITTDKIEINLGFVDVDTSTKERLEQJASK--DKPSTLAV 266
Db      175 YGESGAKYAN-KOLEITSEKSNFSFSDVDIG-----RTASKMSKGDIVHIF 221
QY      267 AGPLTEI-FSDIINOND-RYLIGVDTQSLVYTKRKFFTSILKNLGYSVFSLDLY 324
Db      222 AAGLAGIVBAKMLGSGYVYIGADQDS--YIAPKN-FITSYVKNIGDALYITGE-Y 277
QY      325 TKKSNSRNLAFERGKKSATVYLGIDRFVADIADTSLEGNDKILATEAISAKKEFEK 384
Db      278 IKNNV-----WEGK--VVQMGIRDGVIGLIPNAN-----EF-EYI 310
QY      385 KTIPEAEVKTLEIP 399
Db      311 KVLERKIYVKEIYV 325

```

```

RESULT 3
US-08-396-957A-4
: Sequence 4, Application US/08396957A
: Patent No. 5780041
: GENERAL INFORMATION:
: APPLICANT: SIMPSON, WARREN; SCHWAN, TOM G.
: TITLE OF INVENTION: ANTIGENIC PROTEINS AND
: TITLE OF INVENTION: GENES ENCODING SAME OF BORRELIA BURGDORFERI.
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
: STREET: 345 PARK AVENUE
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: USA
: ZIP: 10154
: COMPUTER READABLE FORM:

```

```

? MEDIUM TYPE: FLOPPY DISK
? COMPUTER: IBM PC COMPATIBLE
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: WORDPERFECT 5.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/396,957A
? FILING DATE: 01-MAR-1995
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/020,245
? FILING DATE: 19-FEB-1993
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/664,731
? FILING DATE: 05-MAY-1991
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/487,716
? FILING DATE: 05-MAY-1990
? ATTORNEY/AGENT INFORMATION:
? NAME: DOROTHY R. AUTH
? REGISTRATION NUMBER: 36,434
? REFERENCE/DOCKET NUMBER: 2026-4018US4
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 758-4800
? TELEFAX: (212) 751-6849
? TELEX: 421792
? INFORMATION FOR SEO ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 339
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? HYPOTHETICAL: NO
? ORIGINAL SOURCE:
? ORGANISM: Borrelia burgdorferi
? STRAIN: Sh-2-82
? INDIVIDUAL ISOLATE:
? DEVELOPMENTAL STAGE:
? HAPLOTYPE:
? TISSUE TYPE:
? CELL TYPE:
? CELL LINE:
? ORGANELLER:
? FEATURE:
? NAME/KEY: p39\
? LOCATION:
? IDENTIFICATION METHOD:
? OTHER INFORMATION: p39\ protein
? OTHER INFORMATION: sequence.
? 08-396-957A-4

Query Match 7.5%; Score 161; DB 1; Length 339;
Best Local Similarity 22.8%; Pred. No. 4.5e-07;
Matches 85; Conservative 68; Mismatches 136; Indels 84; Gaps 18;

```

```

? 287 IGVDTDOSLVYTKKFTSLKNGSVFVSLDLYTKKSNRLAGEFGKKSATYV 346
? 239 IGVDEDAVL---APDVITSTKDVGRAL-----NIFT--SHMLTNFEGGK---LIN 285
? 347 LGIKDRFVDIADTSLEGNKKLATLTAISEAKKEFEKTKTIPAEVAKTLEIEMDKOP 406
? 286 YGLKEGVGVFV-----RNPWMISF-----ELEKEIDMLSKKIINKLEIIV----- 325
? 407 DKQOESLDKLTID 419
? 326 -SMKESEKFLKE 337

```

```

RESULT 4
US-08-313-412-1
; Sequence 1, Application US/08313412
; Patent No. 6248583
; GENERAL INFORMATION:
; APPLICANT: Aron Ph.D., Lieselotte
; APPLICANT: Cabello M.D., Felipe
; APPLICANT: Godfrey M.D., Henry P.
; APPLICANT: Schwartz Ph.D., Ira
; TITLE OF INVENTION: Chromosomally-Encoded Membrane
; TITLE OF INVENTION: Protein of Borrelia burgdorferi
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,412
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 35553/1020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1304
; TELEFAX: 716-263-1600
; INFORMATION FOR SEO ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-313-412-1

```

```

Query Match 6.6%; Score 143; DB 4; Length 353;
Best Local Similarity 23.4%; Pred. No. 2.4e-05;
Matches 83; Conservative 54; Mismatches 131; Indels 86; Gaps 18;

? 55 VTAAGTVNDSPNOSGWEALIOGL-----ALTGEITSVDSSTAELGGY 99
? 34 VLAHSFYDGYNSVDGYVKKLRDNFGIKLITKSLRPYIEGRKRLTYDE--AMTEDAY 91
? 100 SSLANTKKNWVLSGFGHGAFTRWMLKIPENKOLFTEKNIILICI-----DWTDTENVIP 155
? 92 EVQKNPLMLFMLI-GYRFSDL-----SVKLSYERPDYIYGIIADAFYGDIO--VPK 139
? 156 GRINLTYTEDEAGWGLYANASTFLAKFPSDPKRSALYIG---GGISPAVYDFIAGYL 212
? 140 NS-LAIKFRNEEAFIAGYIAAKMSRKE-----KIGFLTPGWSHLADFKRGFK 187

```

OY 213 AGIKAMN--LKSDDKRTTTTKIEINLGFVDQDSTKERLDIA----SKOPSPRLVAV 266  
 Db 188 AGIFIANPRKLRLVSKKAPSLFD-----KKGAMALFMKREDKRVYIFPI 232  
 OY 267 ACPLEIEFSDIANONDR-YLIGVDTDOSLVYTKRKNEFTSLKNLGYSVFSLDLYT 325  
 Db 233 AGITGLGVADAKELGPKYVYIGLNDQSYI---APQVITSIHKIGVIVISSEYI- 288  
 OY 326 KKSNNRLAGFEFGKKSATVYLGINDRFPIA-DTSLBNODKKLATAISEAKR 378  
 Db 289 ---NNRVKRG-----GIIDRGLEGVLEIYKDPDVANN--RLVDEVIDLENK 331  
 RESULT 5  
 US-09-336-447A-15  
 ; Sequence 15, Application US/09336447A  
 ; Patent No. 6310190  
 GENERAL INFORMATION:  
 APPLICANT: HANSEN, ERIC J.  
 APPLICANT: AEBI, CHRISTOPH  
 APPLICANT: CORE, LESLIE D.  
 APPLICANT: MACIVER, ISOBEL  
 APPLICANT: FISKE, MICHAEL J.  
 APPLICANT: FRDENBURG, ROSS A.  
 TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS  
 FILE REFERENCE: AMCY:024  
 CURRENT APPLICATION NUMBER: US/09/336,447A  
 NUMBER OF FILING DATE: 1999-06-21  
 CURRENT FILING DATE: 1999-06-21  
 NUMBER OF SEQ ID NOS: 98  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 15  
 LENGTH: 889  
 TYPE: PRF  
 ORGANISM: Moraxella catarrhalis  
 US-09-336-447A-15  
 Query Match 6.4%; Score 137.5; DB 4; Length 889;  
 Best Local Similarity 22.8%; Pred. No. 0.00036;  
 Matches 112; Conservative 78; Mismatches 198; Indels 103; Gaps 22  
 OY 16 ONKQITDVSKISGLVN---ERKSEIIMAAKADANKHFLMALVTAAGTVNDNSFNOS--- 69  
 Db 210 QNEASGDSSTVSGGNNLAEKSSAIG-GEFNLALGNMATT--SGGRQNEASGRSTVA 266  
 70 GMEALQOUG---ALMG--EITSVSSRA-----ELBEKYSLSLA----- 103  
 Db 267 GGEONQALGKYSTVSGQNEASGDRSTVAGGEONQALGKYSTVSGCYRNQATGKGSFAA 326  
 OY 104 ---NTNKNVWVLSGFQHGDAFTRMUKTIPENKOL-FTEKNITILIGDWTDENVJPTGRY 158  
 Db 327 GIDKNAMNDMNVALGNKNTIEGENSVALGSSNTTVKKNQKNVITLISN-TDITDA-QSGSV 384  
 OY 159 I---NLVYKTEBA-----GWLGYANASFLAKKPSDPTRKSAIVITGG-ISPFA 203  
 Db 385 LIGDNTSGKATTAVEDATVVDGLSTLGFAGVSKANSQSTVSGSEGEKROQIVHAGAGRISND 444  
 OY 204 VTDFAGVLGAKIMNLKNSDKKTIITTDKLEINLGFVODQSTKERLEQIASKDKPSTL 263  
 Db 445 STDANQSOLYALAAVADNDQYEDIEKNODDIANKNADIAKNADIAKNADIOITLNDVYKRE---L 500  
 OY 264 LAVAGPLIEFISDIANONDRYLIGVDTDOSLVYTKRKNEFTSLKNLGYSVFSLDLYT 323  
 Db 501 LNLGRLIDQAKADIDONNINHIELAQQODQSHSDIKT-----LKNVNEBLLLELSGHL 553  
 OY 324 YTKKSN--SRNLAGEFEGKKSATVYLG--TKDREVDIADTSLSEGNDKKLA----- 369  
 Db 554 IDOKADLTFRDKALESNVEEGLDLDSGRLLDQKADIAQN--QANIDLAAYNELDQYAO 611  
 OY 370 ---TEAISAKKPEFEKTKTIPAEVRKTLLEIPEMDDKDPDKQOESLDR----- 415  
 Db 612 KQTEALIDALNKASSTNTONTI--EDILAAVNELODAYAKQOTEALIDALNKASSTNTONIAKN 669

```

QY      416  ---LITDINNL 423
          :      |||:
Db      670  QADIANNINI 680

RESULT  6
US-08-973-462-8
; Sequence 8, Application US/08973462B
; Patent No. 6191270
;
GENERAL INFORMATION:
; APPLICANT: DRULHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
; US-08-973-462-8

```

```

Query Match          6.2%: Score 133.5: DB 4: Length 1786;
Best Local Similarity 19.3%: Pred. No. 0.0027;
Matches 85: Conservative 77: Mismatches 145: Indels 133: Gaps 19;

Oy 9 EKSADNONKOITVDYSKISGLVNERKSEI-----MAAKADANKHFGLMALVITAG 59
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 908 ENFOSSEEEKKEIVIV-----IEEKEEVATTLIEVBOAEKRSAN-----TITEIFE 954

Oy 60 TVYNSNF--NOSGWEATQOLGALGTGETTSDSSAELEGKYSLSLANKKNVWLSGROH 117
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 955 NLEENVAESNENVAENLEKLETVFNVLDRVEETVELTSGE--SLENNEM----- 1002

Oy 118 GDAFTWRLKPIENKOLFETERNIIIGIDPTDENTENVIPTGRY---INTYKTEAGWLAG 173
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1003 -----DKAFSEIFDNKKGIO-----ENLL-TGMRSTISYIYOSSE----- 1039

Oy 174 YANASPLAKKFPSPDPTKRSATVIGGGISPAVDTFIAGYLAGI--KAMWIKNSDKTKTIT 231
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1040 -----KYDLNENYVSSILDNIENKKEGLNKLNENISSTEGVQETVT 1080

Oy 232 DKIEINLGFVDYDPTSTKERLEQIASKDQPSLLVAVAGLTETLF---SPITANONDRIYIIG 288
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1081 EHVHONNYVDV-----DVPMKKDQFGLINEAGLGKMFENLEVPKSESQ--VIT 1129

Oy 289 VDTQOSLYVYTKKKNKFFTSILKNLGYSVFSYLSDIYTRKKSNSRNLGAPGFKSATYVLG 348
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1130 VEELKDEPYQKEVEKEVYSIIIEEMEENVDVLEE---EKED----- 1167

Oy 349 IKDRFVDIADTSLD-GNDKRLATEAISAEKKE---FEETKTTIPAEVVRTLTIEPMPD 403
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1168 LTDMKMAIVAESIISDSKETEESIKDKEDVASLVAEEVQDNMDDESVEKYLELKNM-- 1225

Oy 404 KQPKQOESLDKLTIDINNL 423
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1226 -----BEEELMKDAVEINDI 1239

RESULT 7
Sequence 2, Application US/09308375
Patent No. 6300117

```

```

; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Proteases from Gram-Positive Organisms
; FILE REFERENCE: GC94-PCT
; CURRENT APPLICATION NUMBER: US/09/308,375
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: EP9719636.4
; EARLIER FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2285
; TYPE: prt
; ORGANISM: Bacillus subtilis
; US-09-308-375-2

```

```

Query Match          5.5%; Score 119; DB 4; Length 2285;
Best Local Similarity 18.9%; Pred. No. 0.096;
Matches 99; Conservative 82; Mismatches 184; Indels 158; Gaps 22;

```

```

10 EKASDNOKOITDVSKISGIVNERKSEIMAAKADANKHFGNLMAIYTA---GGTVADNSF 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
466 DKINEDNNYAVATTLDLANSIRKAGSTASTFGEVLENDLIGYTAIASTRESGNIVGNSL 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67 -----NOSGEAIOOLG---ALTGGEITVSSTAELEGKYSILANTNK----- 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
526 KITPARIIGNNOSSIKALEOIGISVKTAGGEAKASADLISFVAGKWPILSDAKONTSIGV 585
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
108 -NWVVLGSGF-----OHGDAFTFRLKIPENKOL-----FTEKNT 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
586 AGIYQSRFNAMNNFNSIAONAKTAASTGSAWSEOKYADSLQARVNKLQNNFTE--F 643
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
140 IILGIDMTDENYI---PGRYINL-TYKTEAGW---LAGYANASFLAKFPSPDPK 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
644 AINASAFITSDGLEFTQAAGSLNASTGVKSGFLPPLLAIVSTATILLST--WTRL 701
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
191 RSAIVIG-----GGISPA-----VTDFIAGYLAGIKAMNL 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
702 ASLLIICTRAMOGELTAGLEAGMTRAAVASHVLTALRGILVSTLVGCAFAL--GMAL 760
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
221 K--NSDKTKITTDKIEINLGFVDVDTSTKERLEQIASKDKPSTLLAAGPLTEIFSDI 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
761 ESLISFAEAKKAKDDPEQSOQTNEAITTNK-----DSTDK-----LIQYKEL 805
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
278 IANONRYILGVPTDGLVYTKTKNKFPSILKNLGYSVSVSLDLYTKSNSRNLAGFE 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
806 QYKESRSILTSDEQRYLQVTOQLAOTFPALVGYDSQGNAIL-----KTNKELEKAT 859
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
338 FGKKSATVYIGIDRFVADIADTSLEGNDKLATEAISEAKKEFEKTKTI--PAEEVRYK 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
860 NTKS---YIALKKQ-----ETRDSAKKTEDAKSKELKSKDELKQY 897
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
396 LEIPEMPDK-----OPDKQESLDRKLITDINN 422
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
898 KQIADYNDRGPRKMDLIADDDYKVAADKAKQGMLEAKQSDIES 940
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 8
US-08-409-995-2
; Sequence 2, Application US/08409995
; Patent No. 5646259
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen I.
; APPLICANT: St. Gene III, Joseph W.
; TITLE OF INVENTION: Haemophilus Adhesion Proteins
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albrighton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA

```

```

; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/409,995
; FILING DATE: 24-MAR-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-61053/REF
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; US-08-409-995-2

```

```

Query Match          5.5%; Score 117.5; DB 1; Length 1098;
Best Local Similarity 19.1%; Pred. No. 0.04;
Matches 112; Conservative 75; Mismatches 170; Indels 229; Gaps 28;

```

```

7 TREKESADNOKOITDVSKISG-----LVNERKSEIMAAKA--DANKHGLN- 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
289 TYEFLSADPTETTYVYDSKENGKRTVEKIGAKTSVIEKHKGLFTKANKETKVVGCANA 348
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
52 -----MAIVTAGVYVNDSEFNOGGEAIOQLGALTGGEITSVDSST---AELEGKYS 100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
349 TEDADEKGLVYAKDYI--DAVNKTGWRIKTTDANGQNGDFATVAGSTNTVFASGNCTTA 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
101 SLANTKNNWVVLGSGFGD-----AFTRMRLK-----PENKQ 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
407 TVYNGTGDITVYKIDAKGDLKLDGDKIADDTYALTYVNDGKANNPKQYADVASTDERR 466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
133 LFEKNTII--ILGIDMTDT-----ENVLPGRYI-----NLTYKTEAGW 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
467 LVYAKGLVYALNLSWTTTAAEADGTLDSNABEQEKADKYTEKAGNKLKVKQSGANF 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
171 LAGYANASFLAKFPSPDPKRSALVIGGISPAVTFD----- 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
527 TYSIQDAL-----TGLTSTITLGTGNNKAKTEINKDGLTTPANGAGANNANTISV 576
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
208 -----IAGYLAGIKAW-----NL--KNSDKKT 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
577 TKDGISAGGSVKNVYSGLKKFEDANFDPLTSSADNLTKNDDAYKGLNLDKRGTDKOT 636
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
228 KITTDKIEINLGFVDVDTSTKERLEQIASKDKPSTLLAAGPLTEIFSDIINQND--RYL 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
637 PYVADNTAAIVG-DLRG-----LGWVISADK-----TTGSGSTE-YHDQVRRANNEKFK 662
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
287 IG-----VPTDGLVYTKTK-----NKEF-----TSILKNLGYSVFVSLSL 333
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
683 SGNGINVGKTVNGRREITFELAKGEVYKSNFEFVKEFNKGTSLVK-----VGDK 733
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
324 YTKSNSRNLAG---PEFGKKSATVYIGIDRFVADIADTS-----LEGN----- 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
734 YYSKEDIDLTGGPKLKDGTVAARKYODKGGKVVSVTDNTEAITTKGSGSYVIGNOVADA 793
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
365 -----DKLATEAISEAKKEFEKTKTIPEAEVRYKLEIPEMPDK 404
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
794 IAKSGFELGLADER--DAKRAFDDKTKALSA-----GTTTEIVNADHK 833
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 9

```

US-08-685-467-2  
; Sequence 2, Application US/08685467  
; Patent No. 6060059  
; GENERAL INFORMATION:  
; APPLICANT: St. Gene III, Joseph W.  
; APPLICANT: Barenkamp, Stephen J.  
; TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Flehr, Holbach, Teel, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685.467  
; FILING DATE: 22-JUL-1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/409,995  
; FILING DATE: 24-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Silva, Robin M.  
; REGISTRATION NUMBER: 38,304  
; REFERENCE/DOCKET NUMBER: A-61053-2/RPT/RMS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1098 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
US-08-685-467-2

Query Match 5.5%; Score 117.5; DB 3; Length 1098;  
Best Local Similarity 19.1%; Pred. No. 0.04;  
Matches 112; Conservative 75; Mismatches 170; Indels 229; Gaps 28;  
Db 7 TKEKSADNOKQITDVSKISG-----LYNERKSEIMAAKA--DANKHFGIN- 51  
Db 289 TVEFLSADTETTYVDSKENGKRTVEKIGAKTSVIEKEDKLTFGANKKETNKVDGANA 348  
QY 52 -----MAITAGGVNDNSFNOSGWEAIOQLALTGGEITSVDSST---AELEGKYS 100  
Db 349 TEDADEGGLVTAKDVI--DAVNKTGWRIKTTDANGQGDPAIVASGTVTFASGNGTTA 406  
QY 101 SLANTNKNVWVLSGFQHD-----AFTRWLKI-----PENKQ 132  
Db 407 TVTNGTDGITVYDAKVGDGLKLDGKIAADTTALTVDGNKANNPKGVADVASTDEKK 466  
QY 133 LFTENNI--ILGIDWPT-----ENVIPITGRI-----NLTYTEEGW 170  
Db 467 LVYAKGLVLTALNSLWTTTAAEADGGLDGNASDEQVAKGKVFYKAGKNLKVQEGANF 526  
QY 171 LAGYANASFLAKKPPSPDTPKRSIAIVIGGISPAYTDF----- 207  
Db 527 TYSLODAL-----TGLTSTILGNGNGAKTEINKDGLTTPANGAGANNANTISV 576  
QY 208 -----IAGYLAGIKAM-----NL--KNSDKT 227  
Db 577 TKDGISAGGQSVKNVVGSLKKFKGDANFDPITSSADNLTKQNDAYKGLTINDEKGTDKQT 636

QY 228 KITTDKIEINIGFDVQDSTKERLEQIASKDKPSTLLAVAGLTFEIPFDIANDND-RYL 286  
Db 637 PVVADNTAATVG-DLRG-----LGWVISADK-----TTGGSTE-YHDQVRANAEVFK 682  
QY 287 IG-----VDTOOSLVYTKT-----KFF-----TSLKMLGVSFVSVDL 323  
Db 683 SGNINVSCKTVNGRRETFELAKEGVYKSNFEYKETNGKETSIVK-----VQDK 733  
QY 324 YTKKNSRNLG---FERGKSAIVYLGIDREFDIADTS-----LEGN----- 364  
Db 734 YSKREDIDLTTGQPKLKQGNVYAAKYQDKGVYSVTNTENTINKSGVYTGQVADA 793  
QY 365 -----DKKLTAEISAEKKEFEKTKTTPAEVVRKLTLEIPMPDK 404  
Db 794 IAKSGFELGLADEA--DAKRAFDKTKALSA-----GTETIVAHDK 833

RESULT 10  
US-09-377-155-32  
; Sequence 32, Application US/09377155  
; Patent No. 6197312  
; GENERAL INFORMATION:  
; APPLICANT: PEAK, Ian Richard Anselm  
; APPLICANT: JENNINGS, Michael Paul  
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN  
; FILE REFERENCE: 065064/0128  
; CURRENT APPLICATION NUMBER: US/09/377.155  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: PCT/AU98/01031  
; PRIOR FILING DATE: 1998-12-14  
; PRIOR APPLICATION NUMBER: GB 9726398.2  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 32  
; LENGTH: 1098  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
US-09-377-155-32

Query Match 5.5%; Score 117.5; DB 4; Length 1098;  
Best Local Similarity 19.1%; Pred. No. 0.04;  
Matches 112; Conservative 75; Mismatches 170; Indels 229; Gaps 28;  
Db 7 TKEKSADNOKQITDVSKISG-----LYNERKSEIMAAKA--DANKHFGIN- 51  
Db 289 TVEFLSADTETTYVDSKENGKRTVEKIGAKTSVIEKEDKLTFGANKKETNKVDGANA 348  
QY 52 -----MAITAGGVNDNSFNOSGWEAIOQLALTGGEITSVDSST---AELEGKYS 100  
Db 349 TEDADEGGLVTAKDVI--DAVNKTGWRIKTTDANGQGDPAIVASGTVTFASGNGTTA 406  
QY 101 SLANTNKNVWVLSGFQHD-----AFTRWLKI-----PENKQ 132  
Db 407 TVTNGTDGITVYDAKVGDGLKLDGKIAADTTALTVDGNKANNPKGVADVASTDEKK 466  
QY 133 LFTENNI--ILGIDWPT-----ENVIPITGRI-----NLTYTEEGW 170  
Db 467 LVYAKGLVLTALNSLWTTTAAEADGGLDGNASDEQVAKGKVFYKAGKNLKVQEGANF 526  
QY 171 LAGYANASFLAKKPPSPDTPKRSIAIVIGGISPAYTDF----- 207  
Db 527 TYSLODAL-----TGLTSTILGNGNGAKTEINKDGLTTPANGAGANNANTISV 576  
QY 208 -----IAGYLAGIKAM-----NL--KNSDKT 227  
Db 577 TKDGISAGGQSVKNVVGSLKKFKGDANFDPITSSADNLTKQNDAYKGLTINDEKGTDKQT 636  
QY 228 KITTDKIEINIGFDVQDSTKERLEQIASKDKPSTLLAVAGLTFEIPFDIANDND-RYL 286  
Db 637 PVVADNTAATVG-DLRG-----LGWVISADK-----TTGGSTE-YHDQVRANAEVFK 682



QY 287 IG-----VDTDSLVYTKT-----NKEF-----TSILNIGYVFSVLSL 323  
 DB 683 SGNGINVSCTVNGRREITFELAKGEVKSNETVAKETNGKETSIVK-----VGDK 733  
 QY 324 YTKKSNSRLAG---PEFGKKSATVYLGIDRFVDIADTS-----LEGN----- 364  
 DB 734 YSKEDIDLTGQPKLKDNTVAKYODKGGKVVSVTDNTAEATITNGSGYVGNVADA 793  
 QY 365 -----DKLATEAISEAKKEFEKTKTIPAEVKTLEPEMDK 404  
 DB 794 IAKSGFELGLADEA--DAKRAFDDKTKALSA-----GTEIVNADHK 833

RESULT 11  
 US-08-913-942-2  
 ; Sequence 2, Application US/08913942  
 ; Patent No. 6200578  
 ; GENERAL INFORMATION:  
 ; APPLICANT: St. Geme, Joseph  
 ; APPLICANT: Barenkamp, Stephen J.  
 ; TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP  
 ; STREET: Four Embarcadero Center, Suite 3400  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: United States  
 ; ZIP: 94111-4187  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/913,942  
 ; FILING DATE: 29-DEC-1997  
 ; CLASSIFICATION: 514  
 ; APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/409,995  
 ; FILING DATE: 24-MAR-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US96/4031  
 ; FILING DATE: 22-MAR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Vance, Dolly A.  
 ; REGISTRATION NUMBER: 39,054  
 ; REFERENCE/DOCKET NUMBER: A-61053-1/RFT/RMS/DVA  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 781-1989  
 ; TELEFAX: (415) 398-3249  
 ; TELEX: 910 277299  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1098 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; US-08-913-942-2

Query Match 5.5%; Score 117.5; DB 4; Length 1098;  
 Best Local Similarity 19.1%; Pred. No. 0.04; Indels 229; Gaps 28;  
 Matches 112; Conservative 75; Mismatches 170;

QY 7 TKEKSDNOKOITDVSKISG-----LVNERKSEIIMAAKA--DANKHFGIN- 51  
 DB 289 TVEFLSADTETTVYDVKSKNGKRETEVIGAKTSVIEKDKGLFTGANKNETKNVDGANA 348  
 QY 52 -----MAIVTAGYVNDNSFNOSGWEAIOQLGALTGETTSVDSST-----ALEBKYS 100  
 DB 349 TEDADEGKGLTAVDVI--DAVNKTGMRIKTTDANGONGDFATVAGSQTNTVFASGNGTGA 406

DB 349 TEDADEGKGLTAVDVI--DAVNKTGMRIKTTDANGONGDFATVAGSQTNTVFASGNGTGA 406  
 QY 101 SLANTNKNVWVLSGFOHGD-----AFTBMLKI-----PENKO 132  
 DB 407 TVYNGIDGITVYKDAKVDGDLKIDGKIADDTALTALYNDGKNANNPKGKAVDAVSDEKK 466  
 QY 133 LFTENKII--ILGIDWTD-----ENVIPGRYI-----NLTYKTEBAGW 170  
 DB 467 LVYAKGLVTLNLSMTTAAEDGGTLGDGNASEOEYKAKADKVTFRAGKNLKVKEGANGF 526  
 QY 171 LAGYANASFLAKKFPSPDPTRSALVIGGIGSPAVTDF----- 207  
 DB 527 TVSLADAL-----TGLSTLTGTGNNGAKTEINKDGLTTPANGAGANNANTISV 576  
 QY 208 -----IAGYLAGIKAW-----NL--KNSDKKT 227  
 DB 577 TKDGISAGGOSVKNVYSGLKFFGDANDPLTSSADNLTGNDAYKGLTULDEKIDKOT 636  
 QY 228 KITTDKIEINLGFVDVDTSTKERLEQJASKDKSTLLAVAGPLTEIFSDIIANQND-RYL 286  
 DB 637 PVVADNTAAATVG-DLRG-----LGWVTSADK-----TTGGSFE-YHDQVRANAEYKFK 682  
 QY 287 IG-----VDTDSLVYTKT-----NKEF-----TSILNIGYVFSVLSL 323  
 DB 683 SGNGINVSCTVNGRREITFELAKGEVKSNETVAKETNGKETSIVK-----VGDK 733  
 QY 324 YTKKSNSRLAG---PEFGKKSATVYLGIDRFVDIADTS-----LEGN----- 364  
 DB 734 YSKEDIDLTGQPKLKDNTVAKYODKGGKVVSVTDNTAEATITNGSGYVGNVADA 793  
 QY 365 -----DKLATEAISEAKKEFEKTKTIPAEVKTLEPEMDK 404  
 DB 794 IAKSGFELGLADEA--DAKRAFDDKTKALSA-----GTEIVNADHK 833

RESULT 12  
 US-09-669-974-32  
 ; Sequence 32, Application US/09669974  
 ; Patent No. 6331173  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PEAK, Ian Richard Anselm  
 ; APPLICANT: JENNINGS, Michael Paul  
 ; APPLICANT: MOXON, E. Richard  
 ; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN  
 ; FILE REFERENCE: 065064/0128  
 ; CURRENT APPLICATION NUMBER: US/09/669,974  
 ; CURRENT FILING DATE: 2000-09-26  
 ; PRIOR APPLICATION NUMBER: US 09/377,155  
 ; PRIOR FILING DATE: 1999-08-19  
 ; PRIOR APPLICATION NUMBER: PCT/AU98/01031  
 ; PRIOR FILING DATE: 1998-12-14  
 ; PRIOR APPLICATION NUMBER: GB 9726398.2  
 ; PRIOR FILING DATE: 1997-12-12  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 32  
 ; LENGTH: 1098  
 ; TYPE: PRT  
 ; ORGANISM: Haemophilus influenzae  
 ; US-09-669-974-32

Query Match 5.5%; Score 117.5; DB 4; Length 1098;  
 Best Local Similarity 19.1%; Pred. No. 0.04; Indels 229; Gaps 28;  
 Matches 112; Conservative 75; Mismatches 170;

QY 7 TKEKSDNOKOITDVSKISG-----LVNERKSEIIMAAKA--DANKHFGIN- 51  
 DB 289 TVEFLSADTETTVYDVKSKNGKRETEVIGAKTSVIEKDKGLFTGANKNETKNVDGANA 348  
 QY 52 -----MAIVTAGYVNDNSFNOSGWEAIOQLGALTGETTSVDSST-----ALEBKYS 100  
 DB 349 TEDADEGKGLTAVDVI--DAVNKTGMRIKTTDANGONGDFATVAGSQTNTVFASGNGTGA 406

```

OY 101 SLANTNNKVVWVLSGFOHGD-----AFTRWLKI-----PENKO 132
DB 407 TYTNGTDGTVAYDAKVGGLDGLDDKIAADFTALTVDNGKNANPNKGVADAVASTDEKK 466
OY 133 LFTENNIIT--IIGIDMTDT-----ENVIPGRYI-----NLTYKTEEAGW 170
DB 467 LVTAGGLVYALNSLSMTTAAEADGTLTDGNASEQEVKAGDKVTEKAGKNLKVKEGANF 526
OY 171 LAGVANASFLAKFPSPDKRSAYIGGISPAVTDF-----207
DB 527 TYSLODAL-----TGLTSITLGTGNNCAKTEINKDGLTTPANGAGANNANTISV 576
OY 208 -----IAGYLAGIKAN-----NL--KNSDKKT 227
DB 577 TKDGISAGGQSVKNVNSGLKFGDANFDPULTSSADNLTQNDQAVKGLNLDEKGTROT 636
OY 228 KITTKKIEINLGFVDQDSTKERLQJASKDKPSTLLAVAGPLTEFSDIINOND-RYL 286
DB 637 PVVADNTAATVG-DLRG-----LGMVISADK-----TTGGSTE-YHDOVRNANEVKEK 682
OY 287 IG-----VDTQSLVYTKT-----NKEF-----TSILKNLGYSVFVSLDL 323
DB 683 SCNGINVGSKTYNGRREITFEELAKEGVKSNFTVKEKETSIVK-----VGDK 733
OY 324 YTKKNSRNLG--FEGKKSATVYLGKDFVDIADTS-----LEGN-----364
DB 724 YSKEDIDLTTGQPKLKDNTVAARYODKGVSVTDTEATITNKSGGYTGNOVADA 793
OY 365 -----DKKATEAISEAKKEFEKTKTTPAEVVKTLTEIPMPDK 404
DB 794 IAKSGFELGLADEA--DAKRAFDDTKALSA---GTEIIVAHDK 833

```

```

RESULT 13
US-09-268-347-44
Sequence 44, Application US/09268347
Patent No. 6335182
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
FILE REFERENCE: 1038-860
CURRENT APPLICATION NUMBER: US/09/268,347
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 44
LENGTH: 1098
TYPE: PRP
ORGANISM: Haemophilus influenzae
US-09-268-347-44

```

Query Match 5.5%; Score 117.5; DB 4; Length 1098;  
 Best Local Similarity 19.1%; Pred. No. 0.04; Indels 229; Gaps 28;  
 Matches 112; Conservative 75; Mismatches 170;

```

OY 7 TREKESADNONKQITDVSKISG-----LVNERKSEIMAKA--DANKHEGLN- 51
DB 289 TVEFISATETTTTVDKREKCKRTEVAKIGAKTSYIKEDGKLFPGKANKENKVDGANA 348
OY 52 -----MAVTAGTVNDNSFNOSGWEALQOLGALTGETISVDSST---AELEGKYS 100
DB 349 TEDADEGKGLVYAKDVI--DAVNKTGMRIKTTDANGONGDFATVASGNVTPASGNGTAA 406
OY 101 SLANTNNKVVWVLSGFOHGD-----AFTRWLKI-----PENKO 132
DB 407 TYTNGTDGTVAYDAKVGGLDGLDDKIAADFTALTVDNGKNANPNKGVADAVASTDEKK 466
OY 133 LFTENNIIT--IIGIDMTDT-----ENVIPGRYI-----NLTYKTEEAGW 170
DB 467 LVTAGGLVYALNSLSMTTAAEADGTLTDGNASEQEVKAGDKVTEKAGKNLKVKEGANF 526

```

```

OY 171 LAGVANASFLAKFPSPDKRSAYIGGISPAVTDF-----207
DB 527 TYSLODAL-----TGLTSITLGTGNNCAKTEINKDGLTTPANGAGANNANTISV 576
OY 208 -----IAGYLAGIKAN-----NL--KNSDKKT 227
DB 577 TKDGISAGGQSVKNVNSGLKFGDANFDPULTSSADNLTQNDQAVKGLNLDEKGTROT 636
OY 228 KITTKKIEINLGFVDQDSTKERLQJASKDKPSTLLAVAGPLTEFSDIINOND-RYL 286
DB 637 PVVADNTAATVG-DLRG-----LGMVISADK-----TTGGSTE-YHDOVRNANEVKEK 682
OY 287 IG-----VDTQSLVYTKT-----NKEF-----TSILKNLGYSVFVSLDL 323
DB 683 SCNGINVGSKTYNGRREITFEELAKEGVKSNFTVKEKETSIVK-----VGDK 733
OY 324 YTKKNSRNLG--FEGKKSATVYLGKDFVDIADTS-----LEGN-----364
DB 734 YSKEDIDLTTGQPKLKDNTVAARYODKGVSVTDTEATITNKSGGYTGNOVADA 793
OY 365 -----DKKATEAISEAKKEFEKTKTTPAEVVKTLTEIPMPDK 404
DB 794 IAKSGFELGLADEA--DAKRAFDDTKALSA---GTEIIVAHDK 833

```

```

RESULT 14
US-08-923-992A-4
Sequence 4, Application US/08923992A
Patent No. 6280738
GENERAL INFORMATION:
APPLICANT: Tai, Joseph Y.
APPLICANT: Blake, Milan S.
TITLE OF INVENTION: No. 6280738-19a Fc Binding Forms of the Group B
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/923,992A
FILING DATE: 05-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,707
FILING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 1438, 0140001/RWE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1104 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-923-992A-4

```

Query Match 5.4%; Score 116.5; DB 4; Length 1104;  
 Best Local Similarity 17.9%; Pred. No. 0.051;  
 Matches 85; Conservative 78; Mismatches 208; Indels 103; Gaps 13;





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:29:13 ; Search time 56.14 seconds

(without alignments)  
892.310 Million cell updates/sec

Title: US-09-676-249A-2

Perfect score: 2299  
Sequence: 1 MKRKIKWKEFLGLVPLS.....KQPDKQIISDKLTIDINKN 451

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: A.Geneseq\_032802:\*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 2299  | 100.0       | 451    | 22 | AAU01859    |
| 2          | 2109  | 91.7        | 423    | 22 | AAU01860    |
| 3          | 714.5 | 31.1        | 453    | 14 | AAK40856    |
| 4          | 694.5 | 30.2        | 432    | 16 | AAK67582    |
| 5          | 655   | 28.5        | 448    | 10 | AAK93343    |
| 6          | 394.5 | 17.2        | 428    | 20 | AAK05332    |
| 7          | 387.5 | 16.9        | 429    | 18 | AAK22727    |
| 8          | 227   | 9.9         | 351    | 21 | AAK81632    |
| 9          | 203   | 8.8         | 328    | 19 | AAK55066    |
| 10         | 182   | 7.9         | 339    | 20 | AAK00049    |
| 11         | 182   | 7.9         | 361    | 20 | AAK00048    |

|    |       |     |      |    |          |
|----|-------|-----|------|----|----------|
| 12 | 179   | 7.8 | 341  | 19 | AAK61758 |
| 13 | 173   | 7.5 | 166  | 22 | AAU07352 |
| 14 | 172   | 7.5 | 341  | 14 | AAK33280 |
| 15 | 165.5 | 7.2 | 340  | 14 | AAK31013 |
| 16 | 162   | 7.0 | 322  | 20 | AAK19985 |
| 17 | 162   | 7.0 | 339  | 19 | AAK61757 |
| 18 | 162   | 7.0 | 339  | 20 | AAK19984 |
| 19 | 150   | 6.5 | 223  | 21 | AAK91298 |
| 20 | 144   | 6.3 | 353  | 22 | AAE03741 |
| 21 | 141.5 | 6.2 | 889  | 19 | AAK68208 |
| 22 | 131.5 | 5.7 | 1786 | 18 | AAK24790 |
| 23 | 129.5 | 5.6 | 1558 | 18 | AAK18324 |
| 24 | 129   | 5.6 | 441  | 18 | AAK28236 |
| 25 | 124.5 | 5.4 | 1135 | 21 | AAK84460 |
| 26 | 122.5 | 5.3 | 2285 | 20 | AAK98149 |
| 27 | 122.5 | 5.3 | 2346 | 22 | ABB63519 |
| 28 | 122   | 5.3 | 2411 | 21 | AAK23860 |
| 29 | 121.5 | 5.3 | 834  | 21 | AAK93404 |
| 30 | 121.5 | 5.3 | 1099 | 19 | AAK40538 |
| 31 | 120   | 5.2 | 719  | 22 | AAK79241 |
| 32 | 119   | 5.2 | 776  | 22 | ABB52597 |
| 33 | 119   | 5.2 | 892  | 19 | AAK68203 |
| 34 | 117.5 | 5.1 | 1073 | 21 | AAK01837 |
| 35 | 117.5 | 5.1 | 1079 | 21 | AAK01836 |
| 36 | 117   | 5.1 | 2353 | 17 | AAK93393 |
| 37 | 116.5 | 5.1 | 873  | 19 | AAK68207 |
| 38 | 116   | 5.0 | 1300 | 22 | AAU33407 |
| 39 | 116   | 5.0 | 1300 | 22 | AAU35313 |
| 40 | 115.5 | 5.0 | 1164 | 17 | AAK85781 |
| 41 | 115.5 | 5.0 | 1164 | 17 | AAK40537 |
| 42 | 115.5 | 5.0 | 1164 | 21 | AAK84459 |
| 43 | 114.5 | 5.0 | 1166 | 20 | AAK08643 |
| 44 | 114   | 5.0 | 1333 | 22 | AAU35343 |
| 45 | 113.5 | 4.9 | 466  | 16 | AAK75739 |

ALIGNMENTS

RESULT 1

ID AAU01859 standard; Protein; 451 AA.

AC AAU01859;

DT 07-SEP-2001 (first entry)

DE Mycoplasma hyopneumoniae MHP3 antigen.

KW MHP3; antigen; vaccine; enzootic mycoplasma pneumoniae;

KW antibody; immunoassay; immunotherapy; anti-idiotypic antibody.

XX Mycoplasma hyopneumoniae.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

|    |                     |                         |                    |
|----|---------------------|-------------------------|--------------------|
| FT | Misc-difference 7   | Location/Qualifiers     | B. burgdorferi ant |
| FT | Misc-difference 99  | /note= "Encoded by TGA" | Interleukin-X (IL- |
| FT | Misc-difference 138 | /note= "Encoded by TGA" | p39-beta. Borrell  |
| FT | Misc-difference 152 | /note= "Encoded by TGA" | p39-alpha. Borrell |
| FT | Misc-difference 174 | /note= "Encoded by TGA" | B. burgdorferi ant |
| FT | Misc-difference 198 | /note= "Encoded by TGA" | B. burgdorferi ant |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | B. burgdorferi ant |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Group B Streptococ |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Borrelia burgdorfe |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | M. catarrhalis str |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | P. falciparum live |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Plasmodium falcipa |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Amino acid sequenc |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Bacillus subtilis  |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Drosophila melanog |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Haemophilus influe |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Amino acid sequenc |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Mutant C-beta prot |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Escherichia coli p |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | M. catarrhalis str |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Haemophilus influe |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Haemophilus influe |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Haemophilus adhesi |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | M. catarrhalis str |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Enterococcus faeca |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Enterococcus faeca |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Group B Streptococ |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Group B Streptococ |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Amino acid sequenc |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | S. aureus SdrE pro |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | Enterococcus faeca |
| FT | Misc-difference 246 | /note= "Encoded by TGA" | B31 outer surface  |

```
XX 26-SEP-2000; 2000EP-0308421.
PF 29-SEP-1999; 99US-0156602.
XX
XX (Pf12 ) PFIZER PROD INC.
XX
XX King KW, Madura RA, Rosey EL;
XX
XX WPI: 2001-309781/33.
DR N-PSDB; AAS03285.
XX
XX New apoprotein antigens encoded by mhp3 gene from Mycoplasma
PT hyopneumoniae useful as a vaccine for treating or preventing diseases
PT caused by Mycoplasma hyopneumoniae
XX
XX Claim 11; Page 17-18; 38pp; English.
XX
XX The sequence is Mycoplasma hyopneumoniae MHP3 antigen. MHP3 antigen
and its fragments are useful in manufacturing a vaccine for treating or
preventing a disease or disorder in an animal, especially pig, caused
by M. hyopneumoniae infection e.g. enzootic mycoplasma pneumonia.
CC The mhp3-encoded proteins may be used as immunogens to generate
CC antibodies which immunospecifically bind such an immunogen. The
CC antibodies generated against the antigen are useful in diagnostic
CC immunoassays, passive immunotherapy and generation of anti-idiotypic
CC antibodies. Mhp3 proteins may also be used in immunoassays, e.g. to
CC detect or measure in a biological sample from a vaccinated or
CC potentially infected test animal the presence of antibodies to the
CC antigen, and thus to monitor the immune response and/or to diagnose
CC infection of the animal.
XX
XX Sequence 451 AA:
SQ
Query Match 100.0%; Score 2299; DB 22; Length 451;
Best Local Similarity 100.0%; Pred. NO. 5.2e-165;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKKKIRKNNKFLGLVFPPLSAITATISAGCWDKETKEEKSADNQNKOITDVSKISGLVNE 60
DB 1 mkkkiknnkflglvfpplsaisatissagcwkdkctckeekssadnqnkqitdvskisglvne 60
QY 61 RKSEMAAKADANKHFGALMAIVTNGGVNDNSFNQSSWEAIIQALGALTGGEITSVDST 120
DB 61 rkseimaaakadankhfgalmaiavtnggvndnsfnqssweaiiqalgtgaltgseltdvdst 120
DB 121 AELGKYSLSLANTNKNVWVLSGFQHDAPTRWLTIPENKOLFTEKNITIIIGIDMTDTEHY 180
DB 121 aelgkyslsiantnknvwlsgfngdaptftrwltipenkoltfekenitiiigidmtcteny 180
QY 181 IPTGRYINLTYKTEBAGMAGYANASFLAKKRPSPDTRKSAIVIGGISPAVTDEIAGYL 240
DB 181 iptgryinltykteagmlyanassflakkrpsdpdtrksaiviggispaavtdfiagyl 240
QY 241 AGICAKMNLAKNSDKRTITTDKTEINLGFVDVDTSTKERLEQJASDKRSTLLAVAGPLTE 300
DB 241 agikawnlknadkrtittdkteinlglfvddvdtstkerleqjaskdpsllavagplte 300
QY 301 IFSDIANONDRYLIGVDVDSGLVYTTKNNKFTSILKNLGVSFVSLSDLTYTKSNSRN 360
DB 301 ifsdianondrylignvdvdsqlyvttknnkftsilknlgysvsvlsdlytkksnsrn 360
QY 361 LAGFEFGKKSATVYLGIDRFVDIADTSLGENDKRLATEAISAAKKEDEKTTIPAEV 420
DB 361 lagfegfkksatvylgidrfvdiadtalegndkrlateaisaakkeetktipaeav 420
QY 421 RKTLEIFPMPPKQPPKOESLDKLTIDINKN 451
DB 421 rktleipmpkppkqppkqesldkltidlnkn 451
RESULT 2
```

```
AAU01860
ID AAU01860 standard; Protein; 423 AA.
XX
XX AAU01860;
AC
XX 07-SEP-2001 (first entry)
XX
XX Mycoplasma hyopneumoniae recombinant MHP3 antigen.
DE
XX
XX MHP3: antigen; vaccine; enzootic mycoplasma pneumonia; mutant;
XX antibody; immunoassay; immunotherapy; anti-idiotypic antibody; muclein.
XX
XX Mycoplasma hyopneumoniae.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 70
FT Misc-difference 417..422 /note="Encoded by RCA"
FT Misc-difference 417..422 /note="Encoded by AATWACCATAT"
XX
XX EPI090995-A2.
XX
XX 11-APR-2001.
XX
XX 26-SEP-2000; 2000EP-0308421.
XX
XX 29-SEP-1999; 99US-0156602.
XX
XX (Pf12 ) PFIZER PROD INC.
XX
XX King KW, Madura RA, Rosey EL;
XX
XX WPI: 2001-309781/33.
DR N-PSDB; AAS03286.
XX
XX New apoprotein antigens encoded by mhp3 gene from Mycoplasma
PT hyopneumoniae useful as a vaccine for treating or preventing diseases
PT caused by Mycoplasma hyopneumoniae
XX
XX Claim 4; Page 19-21; 38pp; English.
XX
XX The sequence represents Mycoplasma hyopneumoniae a recombinant MHP3
CC antigen lacking the first 28 amino acids (the putative signal sequence).
CC MHP3 antigen and its fragments are useful in manufacturing a vaccine for
CC treating or preventing a disease or disorder in an animal, especially
CC pig, caused by M. hyopneumoniae infection e.g. enzootic mycoplasma
CC pneumonia. The mhp3-encoded proteins may be used as immunogens to
CC generate antibodies which immunospecifically bind such an immunogen. The
CC antibodies generated against the antigen are useful in diagnostic
CC immunoassays, passive immunotherapy and generation of anti-idiotypic
CC antibodies. Mhp3 proteins may also be used in immunoassays, e.g. to
CC detect or measure in a biological sample from a vaccinated or
CC potentially infected test animal the presence of antibodies to the
CC antigen, and thus to monitor the immune response and/or to diagnose
CC infection of the animal.
XX
XX Sequence 423 AA:
SQ
Query Match 91.7%; Score 2109; DB 22; Length 423;
Best Local Similarity 99.8%; Pred. NO. 9.9e-151;
Matches 414; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 30 WDKETKEEKSADNQNKOITDVSKISGLVNRKSEIMAKADANKHFGALMAIVTAGTGV 89
DB 2 wdketkeekssadnqnkqitdvskisglvnrkseimaaakadankhfgalmaiavtagtgv 61
QY 90 NDNSFNQSSWEAIIQALGALTGGEITSVDSTPAELGKYSLSLANTNKNVWVLSGFQHDAP 149
DB 62 ndnsfnqssweaiiqalgtgaltgseltdvstpaelgkyslsiantnknvwlsgfngdaf 121
QY 150 TRWLTIPENKOLFTEKNITIIIGIDMTDTEHYIPTGRYINLTYKTEBAGMAGYANASFLA 209
DB 150 trwltipenkoltfekenitiiigidmtctenyiptgryinltyktebagnmlyanassfla 209
```

```

Db 122 twlikipenkgjfteknllilglwtdtenvlptrylnllykteangwlagyanasfla 181
Oy 210 KRPSPDTRKSAIVIGGISPATVDFIAGYLAGIKAMNLSKDKRTITTDKIEINLGF 269
Db 182 kfpsdpktrsaiv199gispavtdfiagylagikamlnkskckittdkielelnqfd 241
Oy 270 VODTSTKERLEQIASDKPSTLLAVAGPLTEIFSDIIANONDRYLIGVDTQSLVYTKK 329
Db 242 vqdstckerleqaskdkpscllavagplteifsdliandqndryllqvdtqslvytlck 301
Oy 330 NKFPTSLNLKGSVPSVSLDLYTKSNSRNLAGFEFGKSAVYLGIDRFPDIDTSL 389
Db 302 nfftsilnmlgyvisvtsldlytkksnrlagfegkksatvlylqikdrfvdiadtsl 361
Oy 390 EGNDRKLATEAISEAKKEFEKTKTTPAEVRKLTLEIPEKPDOPKOOESLKL 444
Db 362 egndkklatalealseakkefektktpaevrkltleipekpdopkqgeslkl 416

RESULT 3
AAR40856
AAR40856 standard; Protein: 453 AA.
AAR40856;
XX 07-MAR-1994 (first entry)
XX 43kd regression associated antigen.
DE Regression associated antigen; tumour: immunotherapy;
KW anti-idiotypic antibodies; antibodies; tumour regression.
XX Mycoplasma hyorhinis.
OS
FH Key Location/Qualifiers
FT Misc-difference 80 /note= "Tryptophan encoded by TGA, normal in
FT Misc-difference 124 /note= "Tryptophan encoded by TGA, normal in
FT Misc-difference 124 /note= "Tryptophan encoded by TGA, normal in
FT Misc-difference 165 /note= "Tryptophan encoded by TGA, normal in
FT Misc-difference 165 /note= "Tryptophan encoded by TGA, normal in
FT Misc-difference 344 /note= "Tryptophan encoded by TGA, normal in
FT Misc-difference 344 /note= "Tryptophan encoded by TGA, normal in
PN US5242823-A.
07-SEP-1993.
PF 07-MAR-1986; 86US-0837494.
XX 07-MAR-1986; 86US-0837494.
PR 16-SEP-1987; 87US-0097910.
PR 11-DEC-1987; 87US-0131815.
PR 04-JAN-1988; 88US-0138923.
PR 16-MAR-1990; 90US-0474730.
PR 02-OCT-1992; 92US-0956546.
XX (ITGE-) INT GENETIC ENG INC.
XX Fared GC, Ghosh-dastidar P, Jar-how L, Sen A;
XX WPI: 1993-295229/37.
XX N-PSDB; AAQ47816.
XX DNA encoding a regression-associated antigen from M. hyorhinis -
XX is used to obtain prods. for diagnosis, localisation and therapy
XX of tumours
XX PS Disclosure; Figure 3; 40pp; English.

```

```

XX Regression associated antigens (RAA's) are identified in material
CC from neoplastic cells by their immunological reactivity with
CC regression associated antibodies from the serum of patients
CC diagnosed as undergoing regression of a tumour. RAA's can be used
CC for tumour immunotherapy and for producing and purifying antibodies
CC which can be used for tumour diagnosis, localisation and therapy.
CC The antibodies can also be used for the production of
CC anti-idiotypic antibodies which can also be used in immunotherapy.
XX
SQ Sequence 453 AA:
Query Match 31.1%; Score 714.5; DB 14; Length 453;
Best local similarity 38.2%; Pred. No. 1.3e-45;
Matches 175; Conservative 88; Mismatches 150; Indels 45; Gaps 16;
Oy 14 GLVPLSAIAT-ISAGCWDKETTKEKSADNQNKTITVSKISGLVNERKSEIMAAKADA 72
Db 12 glstvasvalfvscg-----etdkegkll-----rifd-----nsfvkdrqaeleka----- 54
Oy 73 NKEFGIMAIVTAGTYNDSEFNOSSWEAT-----QQLGALTG-GETSVDSSTAELEGKY 127
Db 55 -knidfnlvlltaggtvqgksingslweavlehydqlektlnldrvsgetcnsgelgky 113
Oy 128 SSLANTNKNVWVLSGFQHDGAFTRWLKIPEN-----KQLFTEKNIIILGIDW--TDTEN 179
Db 114 knflngknkvwiltgfgqgefekflqtdsngkkydsllaekkvliavawdskedkd 173
Oy 180 VIPTRYINLTYTEAGWLAGYANASFLAKFPSPDTRKSAIVIGGISPATVDFIAGY 239
Db 174 lkaghnfisllykteegsfagyasckflaykfpndeakrtliapfggghagayctdfagf 233
Oy 240 LAGIKAMNLSKDKRTITTDKIEINLGFVODTSTKERLEQIASDKPSTLLAVAGPLT 299
Db 234 laglakynndpactakvrlsdnmlndtgr-isdnkcatfinglvns--slvlpvagslt 290
Oy 300 EIRSDIIT--ANONDRYLIGVDTQSLVYTKKRFPTSLIKNLGYSVSVSLDLYTKKSN 357
Db 291 ssvvdalksknkdktyllygtdtdgsklfppa-lvfftslshkngrltyeltdlwlkhd 349
Oy 358 SRNLAGEFEK---KSATVYLGIDREFVDIADTSLGNDKRLATEAISEAKKEFEKTK 413
Db 350 skfsgfstrfklnpanatvykgsddivgvsstvadackvkgafneatadfkqld 409
Oy 414 TTPAEVRKLTLEIPEM--PDKOPDKOESLKLITDIN 449
Db 410 anpl-nyksvlgipmlindakdknealneltlkin 446

RESULT 4
AAR67582
AAR67582 standard; Protein: 432 AA.
XX AAR67582;
AC 06-OCT-1995 (first entry)
DT Cancer metastasis protein.
XX Cancer metastasis protein.
DE Cancer metastasis; clinical tests; detection.
KW Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 55 /note= "corresponding codon TGA"
FT Misc-difference 99 /note= "corresponding codon TGA"
FT Misc-difference 140 /note= "corresponding codon TGA"
FT Misc-difference 140 /note= "corresponding codon TGA"
FT Misc-difference 319 /note= "corresponding codon TGA"
FT

```





QY 358 SRNLAGEFEK-----KSATVYLGIDREVDIADPTSLGNDKRLATEBAISEAKKEFEKTK 413  
 Db 349 s-flgsfrsfkltnaanatvykgsd-flgvsnstvadckvkagelnl-atadffkkq 405  
 QY 414 TIPAEVKTLEPEM-PKQPKQOESLDKLTIDN 449  
 Db 406 ampt-nyksvlgjmlindnakheknelikkin 441

RESULT 6  
 ID AAY05332 standard; peptide; 428 AA.  
 AC AAY05332;

DT 25-JUN-1999 (first entry)  
 DE Inflammatory cytokine inducer.  
 KW Inflammatory cytokine inducer; thrombocytopenia; therapy.

XX Synthetic.

PN JP11089582-A.

PD 06-APR-1999.

PF 19-SEP-1997; 97JP-0273649.

PR 19-SEP-1997; 97JP-0273649.

PA (CHUS) CHUGAI PHARM CO LTD.

DR WPI: 1999-281057/24.

DR N-PSDB; AAX33847.

PT New inflammatory cytokine inducer gene and polypeptide - useful for

PS treatment of thrombocytopenia

XX Claim 3; Page 16-17; 22pp; Japanese.

CC This sequence represents the inflammatory cytokine of the

CC invention. The inflammatory cytokine can be used in a drug, which is

CC useful for the treatment of thrombocytopenia.

XX Sequence 428 AA:

Query Match 17.2%; Score 394.5; DB 20; Length 428;  
 Best Local Similarity 26.8%; Pred. No. 1.6e-21;  
 Matches 126; Conservative 101; Mismatches 174; Indels 69; Gaps 20;

QY 5 IKMKNKFLGIVPLSAI-ATISAGCWDKETTKEKSADNQNQITDVSKISGLVNERKS 63  
 Db 1 mkkskkilllgl-splaaalpavavscgnndes-----nlsfkekdskyttanngfq 52

QY 64 EIMAAKADANKHEGLMAITAGTVNDNSFNQSSWEAIOQLGALGETTSVDSSTAEL 123  
 Db 53 vthnael-----lklprvltdegkiddsfngsaftakalnkcgtclnnyeps-snf 106

QY 124 EGYSSSLANTNNKVVWVLSGFQHDAPFTRMKIPENKOLFTEKNIIILGIDWDTENVVPT 183  
 Db 107 esaynasalagnhkivwlnngfkbgqslky--ldahreelerngkijgidf-dleleykw 163

QY 184 GRYINTVYTEAGWLAGYANASFLAKKPPSDTKRSATVIGGISTPATVDFIAGIAGI 243  
 Db 164 --fysqdnikesaftctgyalaswseq--deskrvasfgygafgyvltlnegrfakg 218

QY 244 KAWNLKNSDKTKTI-TTDRTEINILGFDVODTSTKERLEQIAS-----KDRSTLL 292  
 Db 219 lyyngqh--kskskthtspvkidsgf-----tagekmtvymnvls tpadvkvynphvll 271

QY 293 AVAGPLTEIFSDIIANONDRYLIGVDTQSLVYTKTKNKFETSLKNLGSVFSLSDI- 351  
 Db 272 svagpat--fetvrlankqgyvlgvdsdgml--gqkdrilllevlkhikgavetlidl 327  
 QY 352 -----YTKSNSRNLAGFEFGKKSATVYLGIDREVDIADPTSLGNDKRL-ATEAI 401  
 Db 328 lekeegykyrvvdkkdkkshfctq-----kekwyvaeenhfsnteegakinnk 379  
 QY 402 SEAKKEFEKTKTIPAEVKTLEPEMPD-KQPKQOESLDKLTIDN 450  
 Db 380 keaikmfke-----lpedfvkylnsdkalldgnkldhvsrleaisalnx 425

RESULT 7  
 ID AAW22727 standard; protein; 429 AA.  
 AC AAW22727;

DT 26-FEB-1998 (first entry)

DE Membrane protein M161Ag.

KW M161Ag; membrane protein; leukaemia; treatment.

XX Homo sapiens.

OS Key location/Qualifiers

FT Modified-site 121 /note= "selenocysteine"

FT Modified-site 163 /note= "selenocysteine"

FT Modified-site 185 /note= "selenocysteine"

FT Modified-site 348 /note= "selenocysteine"

FT Modified-site 358 /note= "selenocysteine"

PN JP09157295-A.

PD 17-JUN-1997.

PF 05-DEC-1995; 95JP-0344504.

PR 05-DEC-1995; 95JP-0344504.

PA (KAGA-) KAGAKU GIYUTSU SHINKO JIGYODAN.

DR WPI: 1997-369470/34.

DR N-PSDB; AAT75133.

PT Membrane protein M161Ag - useful in development of leukaemia

PS treating agent

XX Claim 1; Fig 1; 6pp; Japanese.

CC This sequence is a membrane protein designated M161Ag. Oligonucleotide  
 CC probes for the cDNA were designed from putative N-terminal M161Ag peptide  
 CC sequences isolated from P39 cells. M161Ag can be used in the development  
 CC of a leukaemia treating agent.

XX Sequence 429 AA:

Query Match 16.9%; Score 387.5; DB 18; Length 429;  
 Best Local Similarity 26.8%; Pred. No. 5.4e-21;  
 Matches 127; Conservative 96; Mismatches 174; Indels 77; Gaps 21;

QY 5 IKMKNKFLGIVPLSAI-ATISAGCWDKETTKEKSADNQNQIT-----DYSKISGLVN 59  
 Db 1 mkkskkilllgl-splaaalpavavsc-----anndesnlsfkekdsktylttna 48



XX WPI; 1998-272224/24.  
DR N-PSDB; AAV27326.  
XX  
XX Nucleic acid encoding antigenic peptide(s) from Streptococcus  
PT pneumoniae - or their epitope-containing fragments, useful in  
PS protective or therapeutic vaccines, and for diagnosis  
PS Claim 11; Page 50; 118pp; English.  
XX  
XX The present sequence represents a protein from Streptococcus pneumoniae.  
CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein  
CC can be useful in vaccines for inducing protective antibodies against  
CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.  
CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid  
CC are used to detect Streptococcus infection (by usual hybridisation or  
CC amplification methods), also for isolating Streptococcus genes or their  
CC allelic variants. The protein can be used similarly to detect specific  
CC antibodies in standard immunoassays, especially for diagnosing or  
CC monitoring infections. Antibodies which bind the protein are used to  
CC detect corresponding antigens, to purify the protein and for passive  
CC immunisation (optionally coupled to a toxin). Vaccines are administered,  
e.g. by injection, orally or through the skin, typically at 0.01-1000  
(especially 10-300) mu g/ml per dose.  
XX  
SQ Sequence 328 AA;  
Query Match 8.8%; Score 203; DB 19; Length 328;  
Best Local Similarity 25.7%; Pred. No. 3e-07;  
Matches 98; Conservative 56; Mismatches 144; Indels 84; Gaps 19;  
XX  
XX 61 RKSEIMAKKANKHFGIANNAIYAGGTVDNSFNOSWEAIOIGA---LTGGEITSVD 117  
DB 3 rrsrnaassdvk-----tkaavtldtgyvdckfngsawegldawgkehnlpgskgyay 58  
XX  
XX 118 SSTAELE-----GKYSILANTNKNVWLSGFOHGAFTRWMLKIPENKQLFTEKNII 168  
DB 59 qstseadyannlgaagsynlffgv-----gfaInna-----vKdaakehtdlnyv 104  
XX  
XX 169 ILGIDWTDENVIPITGRYINLTATKTEBAGWLAGYANASFLAKKPPSOPTRSAIVIGGI 228  
DB 105 llddvldkqknv-----asvltfadesgyIagvaak-----tktkqygtvgy 149  
XX  
XX 229 -SPAVTDFIGYLAGIKAMWLKNSDKRTKITTDKIEINLGFVDOTSTKRELEOIASDK 287  
DB 150 esevlarifeagfkagv-----aavdpkikqvdyag-sfg-daakgylaaagyaad- 201  
XX  
XX 288 PSTLLAVAGPL-TEIFSIDIANDR-----YLIGVDTDSL--VYT---KTKKFFT 334  
DB 202 --lvygvaagtgagvfaeaklnesrnekvwlygvdrgaeagkyltskdgkesnftlv 259  
XX  
XX 335 SILKNIGYSVFSVSDLYTKSSNRNLAGEFEGKKSATVYLIGIDRPVDIADTSLGENDK 394  
DB 260 stlkqygtvkdI-----snkaergefpqgq---viyyslkdkvdlavtnlseegk 308  
XX  
XX 395 KLATEAISKAKEFEKTKTIP 416  
DB 309 k-----avedakaklldgsavkp 326  
XX  
RESULT 10  
AAV00049  
ID AAV00049 standard; Protein; 339 AA.  
XX  
XX AAV00049;  
XX  
XX 20-APR-1999 (first entry)  
XX  
XX Enterococcus faecalis antigenic polypeptide fragment EF021.  
DE  
XX Enterococcus faecalis; infection; vaccine; immune response; diagnosis;  
KM detection; attenuation; antigenic.

XX  
OS Enterococcus faecalis.  
XX  
XX MO9850554-A2.  
XX  
XX 12-NOV-1998.  
XX  
XX 04-MAY-1998; 98WO-US08959.  
XX  
XX 14-NOV-1997; 97US-0066009.  
XX  
XX 06-MAY-1997; 97US-0044031.  
XX  
XX 16-MAY-1997; 97US-0046655.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Bailey C, Choi GH, Hromocky J A, Kunsch CA;  
XX  
XX WPI; 1999-070095/06.  
XX  
XX N-PSDB; AAX20039.  
XX  
XX New isolated Enterococcus faecalis polynucleotides - used to develop  
PT products for the detection of Enterococcus and for use in vaccines  
PT for prevention or attenuation of Enterococcus infection  
XX  
XX Claim 9; Page 101; 301pp; English.  
XX  
XX  
XX The present sequence represents an antigenic polypeptide fragment  
CC isolated from Enterococcus faecalis. The present invention describes  
CC genes, proteins and antigenic polypeptides isolated from E. faecalis.  
CC The proteins can be used in vaccines for preventing or attenuating an  
CC infection caused by a member of the Enterococcus genus in an animal.  
CC They can also be used for detecting Enterococcus antibodies in a sample.  
CC The nucleotide sequences can be used for detecting Enterococcus nucleic  
CC acids. Products from the present invention can also be used for  
CC screening compounds to identify agonists and antagonists of E. faecalis  
CC protein activity.  
XX  
SQ Sequence 339 AA;  
Query Match 7.9%; Score 182; DB 20; Length 339;  
Best Local Similarity 25.4%; Pred. No. 1.2e-05;  
Matches 104; Conservative 51; Mismatches 139; Indels 116; Gaps 20;  
XX  
XX 61 RKSEIMAKKANKHFGIANNAIYAGGTVDNSFNOSWEAIOIGA---LTGGE----- 112  
DB 7 ktaesggkda-----ahsavltldtgyvdckfngsawegldawgkehnlpgskgyay 62  
XX  
XX 113 ITSVDSS--TAELRGYSSILANTNKNVWLSGFOHGAFTRWMLKIPENKQLFTEKNIIIL 170  
DB 63 lgsndaadytlnldqavsakfnlffgi---gyllkda-----Issaadandpntnftvl 112  
XX  
XX 171 GIDWTDENVIPITGRYINLTATKTEBAGWLAGYANASFLAKKPPSOPTRSAIVIGGISP 230  
DB 113 dddgldgkknv-----vsactfdaeaayIagvaak-----nektkuvqfygeeg 157  
XX  
XX 231 AVTD-FTAGYLAGIKAMWLKNSDKRTKITTDKIEINLGFVDOTSTKRELEOIASKRKPS 289  
DB 158 vvdrtfagagfeky-adaakelojkeiltvdt-----kyaasafadpa 196  
XX  
XX 290 TLAVAGPLTEIFSIDI-----ANONDR-YLIGVDTDSL--VYT 326  
DB 197 kgkalaamaayqngvdlflfhasgatggvfgaaklnesgsdkvwlygvdrgdaagky- 255  
XX  
XX 327 KTK-----NKEFTSILKNIGYSVFSVSDLYTKSSNRNLAGEFEGKKSATVYLIGIDRF 361  
DB 256 ktkgkcdnftlvsctlgvgaavqdanraledk-----fpggehlyv-gikdgg 304  
XX  
XX 382 VDIADTSLGENDKRLATEAISKAKEFEKTKTIPAEVTRTLEIPMPD 431  
DB 305 vdltdgyl--ndkt-----keavktakdkvlgdvkvpeke 339



Db 79 sddlnkrngsdlw-lvgymltda--sllvsenpki--sygldpdygddvqipenll 133  
 QY 182 PTERYINLYKTEEAGMLAGY--ANASFLAKFPSPDTPKRSALVIG--GGISPAVTD-FI 236  
 Db 134 a-----vfrveggafilagyaakksfsgk-----lfiggmknigidafr 174  
 QY 237 AGYAGIKAMNLIKNSDKKRTITTDKIEINLCFDPVDTSTKERLEQIASK--DKPSTLLAV 294  
 Db 175 yyesgakyan-kdlelseynsfadvldg-----rtlaskmgsxgidvlf 221  
 QY 295 AGPLTFEI-FSDIANDND-RYLIGVDTDSLVYTKTKNKEFTSILKNLGYSVFSVLSDLX 352  
 Db 222 aaglagiyvaeaknigdygvyigadqdg--ylapkn-fitsvknigdaYlftge-y 277  
 QY 353 TKRSNSRNLAGFEFGKRSATVYLIGIKDRFVDIADTSLEGNCKRLATEAISEAKFEFEKT 412  
 Db 278 lknnyv-----wegsk-----vvgmgldgyvlgldpnan-----ef-eyl 310  
 QY 413 KTIPEAEVKRTLEIP 427  
 Db 311 kvlerkivnkelivp 325  
 QY 13  
 ID AAU07352  
 AC AAU07352 standard; Protein; 166 AA.  
 XX AAU07352;  
 DT 03-JAN-2002 (first entry)  
 DE Interleukin-X (IL-X).  
 XX Interleukin-X; IL-X; autostimulatory factor; Epstein Barr virus; EBV;  
 KM Lymphoblastoid; B cell.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT Misc-difference 150 /note= "Encoded by TGA"  
 FT Misc-difference 160 /note= "Encoded by TGA"  
 FT  
 XX US2001019713-A1.  
 PN 06-SEP-2001.  
 PD 16-JAN-2001; 2001US-0760541.  
 PF 08-AUG-1994; 94US-0287387.  
 XX 11-FEB-1993; 93US-0026132.  
 PA (BLAZ/) BLAZAR B A.  
 PA (WEBB/) WEBB A C.  
 XX Blazar BA, Webb AC;  
 PI  
 DR WPI: 2001-570181/64.  
 DR N-PSDB; AAS11798.  
 XX  
 PT New IL-X polypeptides useful e.g. for enhancing the growth of cells in  
 PT vitro, for growing primary human B cell cultures to facilitate  
 PT proliferation of these cell lines, as therapeutics or prophylactics, as  
 PT molecular weight standards -  
 XX  
 PS Claim 1; Page 7; 16pp; English.  
 CC The invention relates to a novel secreted autostimulatory factor,  
 CC interleukin-X (IL-X), from an Epstein Barr virus (EBV)-carrying  
 CC lymphoblastoid cell line. IL-X proteins can be used to enhance the growth  
 CC of cells in vitro, or to grow primary human B cell cultures to facilitate  
 CC proliferation of these cell lines, as therapeutics or prophylactics, as

CC molecular weight standards, as inert proteins in an assay, or in the  
 CC detection of antibodies that are immunoreactive with IL-X. Nucleotide  
 CC sequences encoding the peptides could be used to produce recombinant  
 CC peptides or they could be used as probes or primers for diagnostic  
 CC and/or analytical polymerase chain reaction (PCR) procedures, or as DNA  
 CC or RNA sizing standards. The present sequence represents the amino acid  
 CC sequence of IL-X.  
 XX  
 SQ Sequence 166 AA;  
 QY 225 GGGISPAVTDPIAGYLAGIKAMNLIKNSDKKRTKI-TTDKIEINLCFDPVDTSTKERLEQIA 283  
 Db 2 ggaafpvtlfnegfakgillynqkh-kssklyhtspvldsgf-----tagekmvtl 54  
 QY 284 S-----KDKPSTLLAVAGPLTFEISDIANDNDRYLIGVDTDSLVYTKRKNKP 333  
 Db 55 nvlsstpadvkynphvllavagpat--fetvrlankgyvlygdadqgmll--qdkdrll 110  
 QY 334 TSLIKNLGYSVFSVLSPL-----YTKRSNRLAGFEFGKRSATVYLIGIKDRFVD 383  
 Db 111 tsvllkhikgavyetlidlilekeegykykpyvdkdkakkwshfgtc-----kekwi 162  
 QY 384 IAD 386  
 Db 163 vae 165  
 RESULT 14  
 ID AAR33280  
 AC AAR33280 standard; Protein; 341 AA.  
 XX AAR33280;  
 DT 17-DEC-2001 (updated)  
 DT 11-MAY-1993 (first entry)  
 DE  
 XX P39-beta.  
 XX Polymerase chain reaction; PCR; primer; open reading frame; ORF; P39;  
 KM antigen; 39-alpha; 39-beta; serum; human; Lyme patient; p39-beta;  
 KM p39-alpha; promoter; -10; -35; ribosomal binding site; RBS.  
 XX  
 OS Borrelia burgdorferi.  
 XX  
 PN USN7664731-N.  
 XX  
 PD 01-DEC-1992.  
 XX  
 PF 05-MAR-1990; 90US-0487716.  
 XX  
 PR 05-MAR-1990; 90US-0487716.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
 XX  
 PT Schwan TG, Simpson WJ;  
 DR WPI: 1993-017799/02.  
 DR N-PSDB; AAQ34672.  
 XX  
 PT New Borrelia burgdorferi protein and DNA isolates - used in  
 PT detection of Lyme borreliosis infection and in prodn. of  
 PT antibodies and vaccines  
 XX  
 PS Disclosure; Page 36-39; 69pp; English.  
 CC The sequences given in AAR331013 and AAR33280 represent the B.  
 CC burgdorferi proteins P39-alpha and P39-beta. The DNA encoding these  
 CC proteins was isolated by polymerase chain reaction (PCR) using the

CC primer sequences given in AA034673-76. This DNA contained two open  
 CC reading frames (ORF). It was therefore concluded that the p39 antigen  
 CC which has been previously described is not one protein but two, 39-alpha  
 CC and 39-beta. The p39 signal appears to be enhanced when both genes  
 CC are present. Gene 1 encodes a 339 amino acid protein with a  
 CC calculated molecular weight of 36,926 kD. The protein encoded by this  
 CC gene which is reactive with serum from human Lyme patients. The ORF  
 CC of gene 2 has been designated p39-beta. This genes' ORF starts 116  
 CC nucleotides downstream of p39-alpha and encodes a protein of 341 amino  
 CC acids (37,506 kD). The promoter located 5' of p39-alpha appears to be  
 CC present with classic -10 and -35 regions whereas p39-beta lacks a  
 CC recognisable promoter sequence. Both genes have putative ribosomal  
 CC binding sites immediately 5' to the start codons and each is terminated  
 CC with a TAA codon. Comparing the gene sequences of p39-alpha and p39-  
 CC beta indicates that these genes have 62% similarity.  
 CC (Note: Revised entry submitted to correct the patent number format of  
 CC US Government-owned NTIS applications to prevent clashes with ongoing US  
 CC granted patent numbers. For further information please visit the Derwent  
 CC web site at [www.derwent.com/dwpl/updates/ntis\\_us.html](http://www.derwent.com/dwpl/updates/ntis_us.html).)

SQ Sequence 341 AA:

Query Match 7.5%; Score 172; DB 14; Length 341;  
 Best Local Similarity 26.7%; Pred. No. 6,8e-05;  
 Matches 100; Conservative 60; Mismatches 123; Indels 92; Gaps 25;

OY 72 ANKHFGIANNATYAGCTVNDNSFNQSSWEAIOOLGALGTGEITSVDSFAELEGRKSS-- 129  
 DB 24 sskkkiklmilv---dgvldkafinsaneallrlkkdfenleevfs--calsgyvsasy 78  
 OY 130 --LANTNKN---VWLSFGHGDATFRWLKTPENKQLFTEKNIT--ILGIDMTDENVI 181  
 DB 79 sldldhkrngsdlw-lygmjtda--sllvssmpki--sygildpdyddqbpemil 133  
 OY 182 PGRVYINTLYKTEEGAGMLGYANA--SFLAKKPPSPDTRSAIVIG--GGISPAVTD-PI 236  
 DB 134 a-----vfrveggafisgyidakkfsfgk-----lglftgmkgynlvadlr 174  
 OY 237 AGYLAGIKAMNLNSDKRTITTDKIEINLGFVDOTSTKERLEQIASK--DKPSTILAV 294  
 DB 175 yyeegakyan-kdellaeysnfsdvdlg-----rtlaekmykgldvlnf 221  
 OY 295 AGPLREI-FSDLIANOND-RYLIGVDTQSLVYTKKNFPSILKNLGSVFSVLSDLY 352  
 DB 222 aeglegigvleaaknlgdygyigadqgs--ylapkn-fltsvknlgdalyiltge-y 277  
 DB 353 TRKSNRNLAGFEFGKRSATVYLGIKDRFVDIADTSLBGNDKKLATETAISEAKKEFEET 412  
 DB 278 lknmv-----weggk-----vvgmgjldgyiglpnan-----el-eyl 310  
 OY 413 KTIPEAVKLTLEIP 427  
 DB 311 kvlerkivnkelivp 325

RESULT 15

AA031013  
 ID AAR31013 standard; Protein, 340 AA.

AA031013;

17-DEC-2001 (updated)

11-MAY-1993 (first entry)

P39-alpha.

Polymerase chain reaction; PCR; primer; open reading frame; ORF; p39;

antigen; 39-alpha; 39-beta; serum; human; Lyme patient; p39-beta;

p39-alpha; promoter; -10; -35; ribosomal binding site; RBS.

Borrelia burgdorferi.

PN USN7664731-N.  
 XX 01-DEC-1992.  
 PD 05-MAR-1990; 90US-0487716.  
 XX 05-MAR-1990; 90US-0487716.  
 PR 05-MAR-1990; 90US-0487716.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
 PI Schwan TG, Simpson WJ;  
 XX WPI: 1993-017799/02.  
 DR N-PSDB: AA034672.  
 PT New Borrelia burgdorferi protein and DNA isolates - used in  
 PT detection of Lyme borreliosis infection and in prodn. of  
 PT antibodies and vaccines  
 XX  
 PS Disclosure: Page 36-39; 69pp; English.

CC The sequences given in AAR31013 and AAR3280 represent the B.  
 CC burgdorferi proteins p39-alpha and p39-beta. The DNA encoding these  
 CC proteins was isolated by polymerase chain reaction (PCR) using the  
 CC primer sequences given in AA034673-76. This DNA contained two open  
 CC reading frames (ORF). It was therefore concluded that the p39 antigen  
 CC which has been previously described is not one protein but two, 39-alpha  
 CC and 39-beta. The p39 signal appears to be enhanced when both genes  
 CC are present. Gene 1 encodes a 339 amino acid protein with a  
 CC calculated molecular weight of 36,926 kD. The protein encoded by this  
 CC gene which is reactive with serum from human Lyme patients. The ORF  
 CC of gene 2 has been designated p39-beta. This genes' ORF starts 116  
 CC nucleotides downstream of p39-alpha and encodes a protein of 341 amino  
 CC acids (37,506 kD). The promoter located 5' of p39-alpha appears to be  
 CC present with classic -10 and -35 regions whereas p39-beta lacks a  
 CC recognisable promoter sequence. Both genes have putative ribosomal  
 CC binding sites immediately 5' to the start codons and each is terminated  
 CC with a TAA codon. Comparing the gene sequences of p39-alpha and p39-  
 CC beta indicates that these genes have 62% similarity.  
 CC (Note: Revised entry submitted to correct the patent number format of  
 CC US Government-owned NTIS applications to prevent clashes with ongoing US  
 CC granted patent numbers. For further information please visit the Derwent  
 CC web site at [www.derwent.com/dwpl/updates/ntis\\_us.html](http://www.derwent.com/dwpl/updates/ntis_us.html).)

SQ Sequence 340 AA:

Query Match 7.2%; Score 165.5; DB 14; Length 340;  
 Best Local Similarity 22.7%; Pred. No. 0.00021;  
 Matches 85; Conservative 70; Mismatches 134; Indels 85; Gaps 19;

OY 87 GTVNDNSFNQSSWEAIOOLGALGTGEITSVDSST---AELEGKSSLANNNKN-VWVLS 141  
 DB 37 glfdskfneeslingvkvkvefkieleysnyslsdleg-----lkdgsgdlwllg 92  
 OY 142 GFQHGD-AFTRWLKIPEKQLFTEKNITIIIGIDMTDENVIPTGRYINLYKTEEGAWLA 200  
 DB 93 fyrfsvdvaakvaalpmdkylal-----ldpysndpjp-anlygmtrraegeafit 142  
 OY 201 GYANASFLAKKFPSPDTRSAIVIGGISPATVD-FINGYLAGIKAMWLKNSDKTKTITT 259  
 DB 143 gylaaekl-----sktqkigfigiegelvdaifryyeagakyan-----kdklkt 188  
 OY 260 DKI-----EINLGFVDQDSTKERLEQIASKDPSTILAVAGPLREIFSDIANO--NDRY 313  
 DB 189 gylgfsfadleagrsvatmystdeid-----lihaaglgigigalevakelegshy 238  
 OY 314 LIGVDTQSLVYTKKNFPSILKNLGSVFSVLSDLYTRKSNRNLAGFEFGKRSATV 373  
 DB 239 illyvedgajl---apdnvltstckdvgral-----nltf--snhltkntleggk---li 285  
 OY 374 YLGIKDRFVDIADTSLBGNDKKLATETAISEAKKEFEETKTIPEAVKLTLEIPMDPKQ 433

Db 286 nyglkegvvgfv-----rnpkmisf-----elekeidnlssklnkeiivp----- 326  
QY 434 PDKQESLDKLTID 447  
:|:|:  
Db 327 --snkesyekflke 338

Search completed: July 18, 2002, 15:32:18  
Job time: 185 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:29:13 ; Search time 33.42 Seconds  
(without alignments)  
1336.715 Million cell updates/sec

Title: US-09-676-249A-2  
Perfect score: 2299  
Sequence: 1 MKKKIKWKEFLGLVLPPLS.....KQPDKQGESLDKLTIDINKN 451

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Minimum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 554.5 | 24.1        | 461    | 2 B90555 | ABC transporter xy  |
| 2          | 223   | 9.7         | 350    | 2 G95097 | lipoprotein [impor  |
| 3          | 219   | 9.5         | 374    | 2 C97965 | conserved hypotet   |
| 4          | 207   | 9.0         | 350    | 2 F86804 | basic membrane pro  |
| 5          | 193.5 | 8.5         | 357    | 2 D96986 | probable lipoprote  |
| 6          | 187.5 | 8.2         | 516    | 2 C82946 | hypothetical prote  |
| 7          | 185.5 | 8.1         | 350    | 1 C70009 | ABC transporter (l  |
| 8          | 182   | 7.9         | 353    | 1 H71340 | membrane lipoprote  |
| 9          | 181.5 | 7.9         | 359    | 1 F72418 | basic membrane pro  |
| 10         | 178   | 7.7         | 341    | 2 E70147 | basic membrane lip  |
| 11         | 174   | 7.6         | 357    | 2 AH1610 | CD4+ T cell-stimul  |
| 12         | 169.5 | 7.4         | 360    | 2 H70147 | basic membrane pro  |
| 13         | 169   | 7.4         | 357    | 2 AD1248 | CD4+ T cell-stimul  |
| 14         | 168.5 | 7.3         | 525    | 2 C82914 | conserved hypotet   |
| 15         | 164   | 7.1         | 524    | 2 D82944 | hypothetical membr  |
| 16         | 162   | 7.0         | 339    | 2 F70147 | basic membrane pro  |
| 17         | 156.5 | 6.8         | 591    | 2 D64204 | membrane lipoprote  |
| 18         | 147   | 6.4         | 353    | 2 G70147 | basic membrane pro  |
| 19         | 143.5 | 6.2         | 384    | 2 I40867 | hypothetical prote  |
| 20         | 141.5 | 6.2         | 539    | 2 D82886 | conserved hypotet   |
| 21         | 141   | 6.1         | 379    | 2 H75318 | membrane lipoprote  |
| 22         | 140.5 | 6.1         | 337    | 2 B97374 | Delinococcus radiod |
| 23         | 140.5 | 6.1         | 337    | 2 AH2591 | membrane lipoprote  |
| 24         | 132   | 5.7         | 547    | 2 E29504 | mercury(II) reduct  |
| 25         | 132   | 5.7         | 657    | 2 S73428 | probable lipoprote  |
| 26         | 130   | 5.7         | 484    | 2 B90524 | hypothetical prote  |
| 27         | 129.5 | 5.6         | 1140   | 2 S73786 | hypothetical prote  |
| 28         | 129.5 | 5.6         | 1558   | 2 B71603 | RESA-H3 antigen pr  |
| 29         | 127.5 | 5.5         | 763    | 2 A82863 | hypothetical prote  |

|    |       |     |      |          |                    |
|----|-------|-----|------|----------|--------------------|
| 30 | 127   | 5.5 | 349  | 2 F84246 | hypothetical prote |
| 31 | 126   | 5.5 | 326  | 2 G95857 | hypothetical prote |
| 32 | 125   | 5.4 | 1635 | 2 A10452 | hemolysin [impor   |
| 33 | 124.5 | 5.4 | 556  | 2 H82301 | peptide ABC transp |
| 34 | 124.5 | 5.4 | 2269 | 2 T28677 | thoptry protein -  |
| 35 | 123.5 | 5.4 | 322  | 2 F84236 | ABC transporter (l |
| 36 | 123.5 | 5.4 | 626  | 2 C25035 | colicin Ia - Esche |
| 37 | 123.5 | 5.4 | 1031 | 2 C81302 | probable type I sl |
| 38 | 123.5 | 5.3 | 1223 | 2 E88451 | protein K10D2.1 (l |
| 39 | 122.5 | 5.3 | 553  | 1 SMEBH1 | flagellar hook-ass |
| 40 | 122.5 | 5.3 | 2285 | 2 AH0640 | flagellar hook-ass |
| 41 | 122.5 | 5.3 | 2285 | 2 T12796 | probable transglyc |
| 42 | 122.5 | 5.3 | 2346 | 2 T13829 | tptr homolog - fru |
| 43 | 121   | 5.3 | 1546 | 2 G90603 | lipoprotein [impor |
| 44 | 119.5 | 5.2 | 350  | 2 F70139 | exported protein ( |
| 45 | 119.5 | 5.2 | 627  | 2 A41609 | dnak-type molecule |

ALIGNMENTS

|   |  |   |
|---|--|---|
| RESULT 1  | B90555   | ABC transporter xylose-binding lipoprotein [imported] - Mycoplasma pulmonis (strain U |
| C:Species:  | Mycoplasma pulmonis  |   |
| C:Date:   | 24-May-2001  | #sequence=revision 24-May-2001 #text-change 03-Aug-2001                               |
| C:Accession:  | B90555   |   |
| R:Champan, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer, |  |   |
| Nucleic Acids Res.  | 29, 2145-2153, 2001  |   |
| A:Title:  | The complete genome sequence of the murine respiratory pathogen Mycoplasma p |   |
| A:Reference number:   | A99512; M0ID:21267165; PMID:11353084   |   |
| A:Accession:  | B90555   |   |
| A:Status:   | preliminary  |   |
| A:Molecule type:  | DNA  |   |
| A:Residues:   | 1-461 <R>  |   |
| A:Cross-references:   | GB:AL445566; PID:g14089760; PIDN:CAC13519.1; GSPDB:GN00153                   |   |
| A:Experimental source:  | strain UAB C11P  |   |
| C:Genetics:   |  |   |
| A:Gene:   | MYPU_3460  |   |
| A:Genetic code:   | SGC3   |   |
| Query Match   | 24.1% ; Score 554.5 ; DB 2 ; Length 461 ;                                    |   |
| Best Local Similarity   | 30.3% ; Pred. No. 9.6e-27 ;  |   |
| Matches 140 ;   | Conservative 88 ; Mismatches 175 ; Indels 59 ; Gaps 13 ;                     |   |
| QY 5 IKMKKFLGLVPLSLATISACQWKEKESKADNONKOT--DVKISGLVNER- 61                          |  |   |
| DB 1 MKLNR--KLFSILPVALLALALPATF-----VSCANPNKTNLSLDSSKRTDLLSOKE 51                   |  |   |
| QY 62 -----KSEIMAKADANKHFLNMAIYAGTVNDNSFNQSSWEAIOGLATGGEI 113                       |  |   |
| DB 52 VTEQKIVENKIKQASLETK-----VLLITADONIDDKFQNGVYSQKTIADFNKAY 106                   |  |   |
| QY 114 TS-----VDSSTAELEGKYSLSANTNNVWVLSGFQHDGAFTRMLKIPENK 159                       |  |   |
| DB 107 KSQKKEAENOHKLDNYNSAVKDEQNYKVALDNGYTWILITGQOQNEIENFENDENNL 166                |  |   |
| QY 160 QLFPEKNIILIGIDWPTEN-VIPTGRYINLYTTEBAGVIAGANSLAKKPPSDPTK 218                  |  |   |
| DB 167 RRFENKRVKITGVMAWADNANSKIPQGSLSLFTETEGAGQWQASADFLGTAKANEAK 226                |  |   |
| QY 219 RSAIVIGGSLPAVTDPIAGYLAGIKAMN--LKNSDKRTITTDKIRINIGFVDQDSTK 276                |  |   |
| DB 227 RAISAFPGGDPAGVTDLNPFBSIRAMNBEAEMANKKVIYSNVLVDLGF-IPNAEK 285                  |  |   |
| QY 277 EREQLNSKDKPSTLLAVAGPLEIEESDIT--ANQDRLLIGVDQSLVYTKKKKFF 333                   |  |   |
| DB 286 EVSNVYETGKSTSLPVAAGPFTGVVAVLRKDTSDREITVGVDTSISFTWDSKRF 345                   |  |   |
| QY 334 TSTIKNLGYSVFSLDLYT-----KKSNSNLNGFEGKKSAAVYVYIGIDRFVDA 385                    |  |   |
| DB 346 TSTVKNIAFVYQTLALLTKDESVTLKRGNDKFLDS---NPKNVLVTKGLSKAFVNT 402                 |  |   |

OY 386 DTSLEGNDRKATEAISEA--KKEFEKTKTTPAEVRKTTLE 425  
DB 403 KSRVKESTIKTQADTSIQKAIKIDKMNPNNSKKEIKENTNDDLE 444

## RESULT 2

695097  
Lipoprotein [Imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001

C:Accession: G95097  
R:RefSeq: G95097  
A:Authors: Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
M.L.; Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.  
Reference number: A95000; MUID:21357209; PMID:11463916

Accession: G95097

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK74976.1; PID:914972319; GSPDB:GN00164; TIGR:SP4

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP0845

Query Match 9.7%; Score 223; DB 2; Length 350;  
Best Local Similarity 24.8%; Pred. No. 1.7e-06;

Matches 110; Conservative .63; Mismatches 148; Indels 122; Gaps 22;

OY 1 MKKKIKMKNFLGLGVPLSATATIS-AGCMDEKTTKEKESADNONKQITDVSKISGLVN 59  
DB 1 MKKKIKMKNFLGLGVPLSATATIS-AGCMDEKTTKEKESADNONKQITDVSKISGLVN 59  
OY 60 EKKSSEIMAAKADANKHFGSLNMAIVTAGTVDNSFNQSSWEAIIQOLGA--LTGGEITSV 116  
DB 40 -----AAIVTDGVDKSFQSDAMEGLQAMGKEHNLKDNNGFTY 79  
OY 117 DSTTAELE-----GKSSSLANTNNVWVLSFGHGDATFRMLKIPENKQLTETKNI 167  
DB 80 POSTSEADYANNLQQAASYNLI--GFLNNA-----VDAKKEHTDLY 125  
OY 168 IILGIDMTDENVIPITGRINLYTEKTEAGWLAGYANASFLAKRPPSPDPTKSAIVIGGG 227  
DB 126 VALDIVINDQKRV-----ASVTFADNESGYLAGYAAK-----TTTKQVGFVGG 170  
OY 228 I-SPAVTDFIAGYLAGIRAMNLSKSDKTKITTDKIEINLGFVDVDTSTKERLEQIASKD 286  
DB 171 IESEVYSRFEAGFKAGV-----ASVDPSIKVQVDYAG-SFG-DAKKGKTIAAQYAAAGAD 223  
OY 287 KPSITLAAVAGPL-TFIFSDIIANQDR-----YLIGVDTQSL-VTT-----KTKNKF 333  
DB 224 ---IYQVAGGTGACGVFAEANSLSNRPENKRWVIGVDRDQEAEGKTSKDGKESNFVL 280  
OY 334 TSILNKLGYSVFSLDYTKKSNRNLAGFEFGKKSATVYLGIDRFVDIADTSLSEGD 393  
DB 281 VSTLQVGTIVDI-----SNKAERGEFPFGQ---VIYSLKDKVDLAVTNLSEGG 329  
OY 394 KKLATEAISEAKKEFEKTKTIP 416  
DB 330 KK---AVEDAKAKIILDSVKVP 348

## RESULT 3

C97965  
conserved hypothetical protein spr0747 [Imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001

C:Accession: C97965  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; B

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.  
Y. P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.

A:Title: Genome of the bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: C97965

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAK99551.1; PID:915458340; GSPDB:GN00174

C:Genetics:

A:Gene: spr0747

Query Match 9.5%; Score 219; DB 2; Length 374;  
Best Local Similarity 24.6%; Pred. No. 3.3e-06;  
Matches 109; Conservative .62; Mismatches 150; Indels 122; Gaps 21;

OY 1 MKKKIKMKNFLGLGVPLSATATIS-AGCMDEKTTKEKESADNONKQITDVSKISGLVN 59  
DB 25 MKKKIKMKNFLGLGVPLSATATIS-AGCMDEKTTKEKESADNONKQITDVSKISGLVN 59  
OY 60 EKKSSEIMAAKADANKHFGSLNMAIVTAGTVDNSFNQSSWEAIIQOLGA--LTGGEITSV 116  
DB 64 -----AAIVTDGVDKSFQSDAMEGLQAMGKEHNLKDNNGFTY 103  
OY 117 DSTTAELE-----GKSSSLANTNNVWVLSFGHGDATFRMLKIPENKQLTETKNI 167  
DB 104 POSTSEADYANNLQQAASYNLI--GFLNNA-----VDAKKEHTDLY 149  
OY 168 IILGIDMTDENVIPITGRINLYTEKTEAGWLAGYANASFLAKRPPSPDPTKSAIVIGGG 227  
DB 150 VALDIVINDQKRV-----ASVTFADNESGYLAGYAAK-----TTTKQVGFVGG 194  
OY 228 I-SPAVTDFIAGYLAGIRAMNLSKSDKTKITTDKIEINLGFVDVDTSTKERLEQIASKD 286  
DB 195 IESEVYSRFEAGFKAGV-----ASVDPSIKVQVDYAG-SFG-DAKKGKTIAAQYAAAGAD 247  
OY 287 KPSITLAAVAGPL-TFIFSDIIANQDR-----YLIGVDTQSL-VTT-----KTKNKF 333  
DB 248 ---IYQVAGGTGACGVFAEANSLSNRPENKRWVIGVDRDQEAEGKTSKDGKESNFVL 304  
OY 334 TSILNKLGYSVFSLDYTKKSNRNLAGFEFGKKSATVYLGIDRFVDIADTSLSEGD 393  
DB 305 VSTLQVGTIVDI-----SNKAERGEFPFGQ---VIYSLKDKVDLAVTNLSEGG 353  
OY 394 KKLATEAISEAKKEFEKTKTIP 416  
DB 354 KK---AVEDAKAKIILDSVKVP 372

## RESULT 4

F86804  
basic membrane protein A [Imported] - Lactococcus lactis subsp. lactis (strain IL1403

C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001

C:Accession: F86804

R:Botolin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malame, K.; Weisenbach, J.; Eh

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: F86804

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 <STO>

A:Cross-references: GB:AE005176; PID:912724428; PIDN:AAK05356.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: bmpA

Query Match 9.0%; Score 207; DB 2; Length 350;

Best Local Similarity 24.2%; Pred. No. 1.7e-05;  
Matches 107; Conservative 62; Mismatches 156; Indels 118; Gaps 20

```

0Y 1 MKIIKKNNKTLGLGVLPPIASAITIS-----AGCMQKETTREKESADNONKQITTVSKISG 56
Db 1 MKK-----YVASSLALASVAVIAGCCRSHDAAGSK----- 32
0Y 57 LVNRKSEIMAAKADANKHFGLMMAIVTAGGVVNDPNSFNOSSWEAIIQALGALTG-----G 111
Db 33 -----AKTD-----LKAALVTEIGGVNDPSFNOSSMEGLQISMGRKNNMLKNGTG 75
0Y 112 EITSVDSSTABEKKYSISLANTKNKNWVLISGPHGDAPFRMLKIDENKOLTEKKITILG 171
Db 76 YTPPQSSASDADYTTNNVSABOQYKLLFGTGFSLDQATSAAK-----NN-----PKSFVIVD 129
0Y 172 IDMTDENIVPTGRXYINLYKTEEAGMILGAYANASFLAKFPSPDPTKSAI-VIGGJISP 230
Db 130 SVIWDQNV-----ASATADNESAVIAGVAAK-----ATIKNKIGFIOGMQSD 174
0Y 231 AVTDFIAGVLAGIKAMNLKNSDKRTIITDKIEINLGPDVDTSPREKLEQIASKDKPST 290
Db 175 VITPEEGEYEGAGASVN-----PDIKDVQYAGSFSFSAAGKGTIAAAMYAGADGV 224
0Y 291 LLAAGSL-TEISDIIA-----NQNDR-VLIGVDTDOSLV--YF-----KTKNKFTSIL 337
Db 225 VYOCAGGAGVIGVSEAKALNSTENADKAWVJGVDDQDQYELKRYSKDKOSNFPVLVSTI 284
0Y 338 KNLIGSVFVSYLSDLTYTKKSNRMLAG--EEFGKSAATVYLGIKDRFYDIADTSLGDNCK 395
Db 285 KEVG-----NVVKIOIADTKTKGCKRPGGVIYTYDLKNGVNLGL-----DSANSEIK----- 330
0Y 396 LATPAISEAKKEFEKTKTTPAE 418
Db 331 ---DAVAKADIIDGKRIYVSK 350

```

RESULT 5  
 D96986  
 Probable lipoprotein, Med/BMP family [Imported] - Clostridium acetobutylicum  
 C:Species: Clostridium acetobutylicum  
 C:Date: 14-Sep-2001 #sequence.revision 14-Sep-2001 #text.change 14-Sep-2001  
 C:Accession: D96986  
 R:Nolling, J.; Berton, G.; Omelchenko, M.V.; Markarova, X.S.; Zeng, Q.; Gibson, R.; Lee,  
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: D96986  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-357 <KMR>  
 A:Cross-references: GB:AE001437; PIDN:AAK78679.1; PID:q15023581; GSPDB:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 C:Genetics:  
 C:Gene: CAC0702

| Query Match           | 8.5%            | Score 195.5   | DB 2       | Length 357 |
|-----------------------|-----------------|---|------------|------------|
| Best Local Similarity | 22.3%           | Pred. No. 8.7e-05   |            |            |
| Matches 100           | Conservative 72 | Mismatches 156  | Indels 121 | Gaps 20    |
| Oy                    | 1               | MKKKKMKNKFLGTGLVPELPISATATISAGCWMDETKTKREKRSADNÖNKQJINDVSKISGLVNE | 60         |            |
|                       |                 |   |            |            |
| Db                    | 2               | IKKKI-----IAILTTVMIVAGLFPAGC-----SSTSSGSGNS                       | 33         |            |
| Oy                    | 61              | RKSEIMAAKADANKHFGIMALVTAAGTYVNDNSFNQSSNEAIQOOLGALTGSGITVSYST      | 120        |            |
|                       |                 |   |            |            |
| Db                    | 34              | K-----DTRK---VKVGLSTDEGGINDKSFNQAGDEGIKKAKEYSVDYKMAIESKK          | 81         |            |
| Oy                    | 121             | A-ELDEGKSSSLANTKKNVWVLSGFÖHGDAF--TRMLKIPENKÖLFTE-----KNIIILG      | 171        |            |
|                       |                 |   |            |            |
| Db                    | 82              | KDDYDPNÖLSLNDNSDILVFGVGYÖMADLDLAIARKYPPKRFALIIDDAYDQPKPNT---      | 137        |            |

```

OY 172 IDMDTENVITPGYINULTKTEBAGLAGIYANAASLAKKFPSPDPYKNAAL-VI0GGISP 230
Db 138 -----MSLVRKE0GSLFM-----VIAGM-----TKTKIGFVGK0P 173
OY 231 AVTDEIAGYLAGIKAWMLKMSDDKTKFTYTDKIEINLNGFV0DPTST-KERLE0IAGSKDRPS 289
Db 174 LINKLGSYIAGANTVN-----PNTTVK-----NTN0YSDPSK0KEVATSLYNGGDI 223
OY 230 TLLAVAGPLTEIESDIIANONDR-----YLIGVDTDSLYTYTKKFKFTSLIKNLGY5VF 345
Db 224 VYHAAGGAGIGVF-DVAKSLRLDDGKDYMAIGVDDKDAALPYRADVILLTSMVKRVDIATY 282
OY 346 SVLSDLYRKNSKNLAGEFEKGKATVYLGIRDFVADIADTSLG0NDKRLATEAISEAK 405
Db 283 NTVDVLYAGKR-----FEGGKVES---FGKLEDEGVAPIS---KKHVPSEVLSLVD 328
OY 406 K-----EPEEKTKTTPAEVVR 421
Db 329 KYKKAIIIDGKIVDPDVA0K0TEKT0DIK 357

```

```

RESULT      6
C82946
hypothetical protein UU012 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence-revision 20-Aug-2000 #text-change 20-Aug-2000
C:Accession: C82946
R:Glass, J.I., Lefkowitz, E.J.: Glass, J.S., Helner, C.R.: Chen, E.Y.: Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A:Reference number: A62870
A:Accession: C82946
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-516 <GLA>
A:Cross-references: GB:AF002100: GB:AF222894: NID:G6898946: PTDN:AAF30417.1: GSPDB:GNP
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: UU012
A:Genetic code: SGC3

```

```

Query Match      8.2%  Score 187.5; DB 2; Length 516;
Best Local Similarity 23.3%  Pred. No. 0.00046;
Matches 117; Conservative 67; Mismatches 167; Indels 151; Gaps 25;

QY      1 MKKKIKMKNFELGLGVFPPLSATATISAGCWDKETHKEEKRSADNQNKKQITDVASKISGLVNE 60
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1 MKKLKKKVVFLFALIGSVFALGIIITVATSTQCKSTL----- 35

QY      61 RKSEIIMAAKADANKHFGLMALVTAGTVNDNSFNQSSWEALIQQLGALTGETTSDSST 120
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      36 -----NYSQFYW-----TSPSTD 49

QY      121 AELEGKYSILANTNNKNWVLSGPOHGDAFTFWKILPENKQ--LEFTEK---NIILL---G 171
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      50 DEGFQTKTKSKMSDGRALLMPGYH-----PELQALVNDKFDPLILALLIDG 99

QY      172 IDMTDTENVIPGT--RYINLTYYKTEAGLAGVYANASFLAK--FPSPDPTKRSATIVIG 226
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      100 YVNNQNNKAEFYGADRVADVVEYKEDAEALFGISIAAAYMLNNSQAVFGAD---NKLTMWG 155

QY      227 --GISPA-VTDSIAGYLAGIKAMNLTKNSDKTK--ITTDKTEINL-----GPD 269
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      156 YVGINAKNTNTNLAQFDLGVKWNANEKLDKNIKQSGTOETKWKIINVEQVYASSSAGGFQ 215

QY      270 VQDTSRKELEQIASQDKRSTLLAAGPPLTEI--FSDIINANODRY-LIGVD---TDSGL 323
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      216 SSSANAKKTIQGLITRG--ADLILPALIPOVGVLAATVEALITTSIHANVGVCADVEINDDAI 274

QY      324 VYTKTKNFFTSIL--KNLGYSVESVL--DLYTKKSNRNIAFGFEKRSATVYIGI- 377
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      275 --NKKTDKINHLGSKNNK--GVIRFSITPRDPAITKILENIGKESLSKSDQIVISGE 331

```

[illegible][illegible]



A:Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MWID:21537279; PMID:11679669  
A:Accession: AH1610  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-357 <GLA>  
A:Cross-references: GB:AL592022; PIRN:CNC96656.1; PID:g16413896; GSPDB:GND0178  
A:Experimental source: strain ClIP11262  
C:Genetics:  
A:Gene: tcsA

|                           |       |                  |            |            |
|---------------------------|-------|------------------|------------|------------|
| Query Match               | 7.68  | Score 174        | DB 2       | Length 357 |
| Best Local Similarity     | 24.28 | Pred. No. 0.0018 |            |            |
| Matches 101, Conservative | 62    | Mismatches 137   | Indels 118 | Gaps 23    |

QY 55 SGEV-----NEKRSITMAAKKADKANKHCLNNAITYTAGTYVNDSPFNOSSEWAIQOLG 106  
 16 SGEVYAGACGSSDDKRS-----GDDKSSKDFYAMVTDITDGVDDRSPFNOSAEGLOKRFG 69  
 QY 107 ALT-----GGEITVSDDSTVALEGRKYSLSLANTNNKNVVL-----GFOGDAFTFMKL 154  
 70 KANMEKCTDDEVNYLQGSSEADYK-----TNLNTAVRSYDYLIGCYIGYIKLKAIEEVS 123

|    |     |   |     |
|----|-----|---|-----|
| QY | 155 | -IPENKOLETEKNNIIIGIDMTDTENNIIPPIGRYINLTYYKTEEGWLAGYANASFLAKKP | 213 |
| Db | 124 | QKPKNOFAIVDDLT-----DSDNNVNSIG-----FKNDSDSYLAGVYAGL-----       | 163 |
| QY | 214 | SDPTKRSALVIGGIGISPAVTD-FLAGYLAGIKAWNNLSDKKRTITDKLEINLGEVQ-    | 271 |
| Db | 164 | --TTKTNVGVGVGVKGVVIDREFAGTAAKA-----VNPNAQIDVQ                 | 205 |

[illegible]

|        |     |   |     |
|--------|-----|---|-----|
| OY     | 373 | VYLGIKORFVADIADTSLGENDKRIATAEIAISEAKKEFEKTKRTIIPAEVEVKLEIPEMP | 430 |
|        | :   | :::   | ::  |
| Db     | 313 | IERYLDRDAVGLESE-----HODNISKDVLAKE-----EYKORIVDGD-----IKVPERP  | 357 |
|        | :   | :::   | ::  |
| RESULT | 12  |   |     |
| 47     |     |   |     |

[illegible]

A:Authors: Smith, H.O.; Venter, J.C  
A:Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*  
A:Reference number: A70100; MUID:9805943  
A:Accession: H70147  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-360 <KEY>  
A:Crossreference:CPDB, E0001144, CP, E0000732, NID, 05689901, NTRY, 0000505

Query Match 7.48; Score 169.5; DB 2; Length 360;  
 Best Local Similarity 25.48; Ptd. No. 0.0035;  
 Matches 88; Conservative 64; Mismatches 137; Indels 57; Gaps

OY 65 IMAKDANHFGLMAYTAGTVDNDSNOSSWEAIOQLGALTGETTSVDSSTAELE 124  
 : : | : : : : | : ||| :|| :| :  
 : : : : | :||| :|| :| :  
 :

```

Db      34 VACSSDDGKSEAKTVLSL-VDGAFDGKGFNESSKAIKRLKADINIIIEKASTGNSYL  92
OY      125 KRYSSILANTNKN-VWVLSGFQHGDAFTFRLKIPKOKOLFTEKNIILIGIDWTDENVNPT  183
Db      93 GDIANLEGGNSLIMGII-GFRLSDIL--FQRASENVSV----NYAI--IEGVDEIQIPK  143
OY      184 GRVILYKTEKTEAGVLAGYANASFLAKKEPSPPTKRSIAIVIGGGSRAVTD-FINGVLAG  242
Db      144 -NLNLSIRSEEVAFIAG-----FASKASTGKIGVYGGVGRGVLESFVMTGFEAG  193
OY      243 IKAMNLSKDKKFKTYTDKI---EINLGFVDQDSTYKERLEQJASKDKPSTLLAVAGPL  298
Db      194 AKYAN-----SNIKVVSQYGVTFEGDFGLG-----RSTASNM---YRGGVDIIFAAAG-L  238
OY      299 TEIFEDIIANO--NDRYVLIGVDQDLSVYTKTKNKFPSILKNGVSSEFVSLDLYTKKS  356
Db      239 SGICVITEAKELGPRHYIIGVDQDQSYL--APNNVIVSAVKKQDLSLWYSL-----TKKY  290
OY      357 NSRNLAGEFEFGKSAFTVYLGIK-DREVDIADTSEGDNKKIATEAI  401
Db      291 LETGV--LDGGR---TWEIIGKEGDLGLVLENENLKSNSSEIYNKSL  331

```

|  |    |
|--|----|
| RESULT   | 13 |
| AD1248   |    |
| CD4+ T cell-stimulating antigen, lipoprotein [Imported] - Listeria monocytogenes (str) |    |
| C;Species: Listeria monocytogenes  |    |
| C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001            |    |
| C;Accession: AD1248  |    |

A. Authors: Kreft, J., Kunh, M., Kunst, F., Kurepkat, G., Madueno, E., Maloum, A., Mok, C., Schluerer, T., Simoes, N., Tlertez, A., Vazquez-Boland, J.A., Voss, H., Wehler, A. Title: Comparative genomics of *Listeria* species. Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: A01246  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-357 <GtAA>  
A/Cross-references: GB:NC\_003210; PIDN:CAC99466.1; PID:g16410817; GSPDB:GN00177  
A/Experimental source: strain EGD-e  
C/Genetics:  
A/Gene: tcsA

```

Query Match          7.4%: Score 169; DB 2; Length 357;
Best Local Similarity 23.8%: Pred. No. 0.0037;
Matches 98; Conservative 63; Mismatches 144; Indels 106; Gaps 22;

54  TSGLVNRRKSEIMAKADANKHFGINMAYTAGGTVDNDSFSSWEIITDGLALT----- 109
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

```

D0 19 ILGAGSSSDKKSSDKSSKDF--TVAMVDTGTGVDDRNSQSGAMEGLQAFGAGNMER 76
QY 110 GGEITSDSSAELEGRKYSLSANTNNKNWVLS-----GFOHGDAFTRLK-IPEKQ 160
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 77 GTDGYNYLQSGASEADYK-----TNLTAVASDYLITGIGYKTLKDAIEEVSQKQPNQF 130
QY 161 LFEETNNIIIGIDMTDPEVNIPTGRYINLTAKYTEEAGMLAGYANASFLAKFFPSDPIKRS 220
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

|    |     |  |     |
|----|-----|--|-----|
| Dd | 131 | AIVDDTL-----DSDNNYSIG-----EKNDGSLVGVNAGL-----TTKTN           | 168 |
| Oy | 221 | AIVGGISPAVTD-PIAGYLAGIKKANLNKNSDKKITTDKIEINLGPDVQ-----DTSTK  | 276 |
|    |     | :     :       :     :  |     |
| Dd | 169 | KVGTVGGKGVIDRFENGFTAGYKA-----VAPNQIDVQVANDPAAKA              | 212 |
| Oy | 277 | ERLEQIISKDKPS-----TLAVALGPLTEIRSDITANONDR-----YLIGVDIQD----- | 321 |

D<sub>b</sub>    213 DKGGQIASMYSSTSGVDVIFHAGGTGNGVFAR-AKNIKKDPSRAVVYIGVDRQMDEGR    271  
Q<sub>y</sub>    322 -SLVYTTRKNKEFFSIILKNLGYSFVSLSLTYYTKSNRNLAG-FEEGSKAATYYLGLIKD    379  
       :    |    ||    :    :    |    |    |    |    :    :

```

Db 272 VTANDKRDYNTLTLSLRPDIA-----EDLAT-----RAKAGDFPGTR--IEYGLDK 319
Qy 380 RFVDIADTSLGNDKRLATEAISAKKEFEKTKTIPAEYVRKLTLEIPEMP 430
    |:::      ::::      |:::      ::::      |
Db 320 DAVGLSE-----HGDNISKDVLAVE---EYKQIVDGD-----IKPEKRP 357

```







```

QY 135 KNWVLSQFHGDQFTRLKLPREKKOLTEKENIIILIGD-WTFQENVIPIGRNYINLYKT 193
Db 93 FDLIYGVGYLMEDELS---ELADQR-----KNTFPAIIDAVYKDNV-----ASTIFKE 138
QY 194 EBAQGLAGIYANASFLAKKFPSPDPKRSYI-VIGGISPAVDTIAGYLAGIKANNLNKSD 253
Db 139 QEGSFLVGVAAAL-----SKSGKIGVGMSEBELKKFPEVGRAGVAONPK-- 186
QY 253 KKTATTDKIEINLGFVDVT--STKERLEQIASKDRPSTLLAVAGPL--TEIFSDIAN- 308
Db 187 -----AVVEVAYAGGFQADYQKATASM---YKSGVDYIYHAGATGIGVFTE-AKNL 236
QY 309 -----QNDRLIGVDTDO---SLVYTTKTKKFFSILKNIGYSVFSVSLDLYTKKSSRN 360
Db 237 KKEDPKRDVWYIGVDKDDYAGQVEGTDDNTLTLSMYK-----KYDVTVEDV-TKKASDGR 291
QY 361 LAGEFEGKRSATVYLKIDREFVDIADTSLSEGNDRKLLATEAISEAKKEF-----EEK 411
Db 292 FPGGP-----TLVYGLDDQGVGISPSQNONSDVYI--KANVDKKKKIIDLGLLEIPATEKE 343
QY 412 TKTIPIAE 418
Db 344 LKTFRAE 350

```

| ID | BMPA_BORGA   | STANDARD: | PRT:                    | 325 AA.                        |
|----|--|-----------|-------------------------|--------------------------------|
| AC | 01357;   | 021360;   |                         |                                |
| DT | 15-DEC-1998  | (Rel. 37, | Created)                |                                |
| DT | 15-DEC-1998  | (Rel. 37, | Last sequence update)   |                                |
| DT | 15-DEC-1998  | (Rel. 37, | Last annotation update) |                                |
| DE | Basic membrane protein A precursor (Immunodominant antigen P39) (Fragment).  |           |                         |                                |
| GN | BMPA.  |           |                         |                                |
| OS | Borrelia garinii.  |           |                         |                                |
| OC | Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.   |           |                         |                                |
| OX | NCBI_TaxID=29519;  |           |                         |                                |
| RN | [1]  |           |                         |                                |
| RP | SEQUENCE FROM N.A.   |           |                         |                                |
| RC | STRAIN=PBI, AND PLT.   |           |                         |                                |
| RX | MEDLINE=98010210; PubMed=9350727;  |           |                         |                                |
| RA | Roesler D., Hauser U., Wilske B.;  |           |                         |                                |
| RT | "Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnosis.";  |           |                         |                                |
| RT | J. Clin. Microbiol. 35:2752-2758(1997).  |           |                         |                                |
| CC | -1- FUNCTION: NOT KNOWN; IMMUNOGENIC PROTEIN.  |           |                         |                                |
| CC | -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).   |           |                         |                                |
| CC | -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.   |           |                         |                                |
| CC | -----  |           |                         |                                |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ). |           |                         |                                |
| CC | -----  |           |                         |                                |
| DR | EMBL; X97244; CA65883.1; -;  |           |                         |                                |
| DR | EMBL; X97238; CA65877.1; -;  |           |                         |                                |
| DR | InterPro; IPR003760; Bmp.  |           |                         |                                |
| DR | InterPro; IPR000437; Prok_Lipoprot.  |           |                         |                                |
| DR | Pfam; PF02608; Bmp; 1.   |           |                         |                                |
| DR | PROSITE; PS00013; PROKAR_LIPOPROTEIN_PARTIAL.  |           |                         |                                |
| KW | Antigen; Membrane; Lipoprotein; Signal.  |           |                         |                                |
| FT | NON_TER  | 1         |                         |                                |
| FT | SIGNAL   | <1        | 3                       | PROBABLE.                      |
| FT | CHAIN  | 4         | 325                     | BASIC MEMBRANE PROTEIN A.      |
| FT | LIPID  | 4         | 4                       | N-ACYL DIGLYCERIDE (PROBABLE). |
| FT | VARIANT  | 8         | 8                       | G -> D (IN STRAIN PL1).        |
| FT | VARIANT  | 12        | 12                      | S -> N (IN STRAIN PL1).        |

|    |          |     |       |                            |
|----|----------|-----|-------|----------------------------|
| FT | VARIANT  | 93  | 93    | P -> S (IN STRAIN PLY).    |
| FT | VARIANT  | 111 | 111   | T -> A (IN STRAIN PLY).    |
| FT | VARIANT  | 180 | 180   | N -> D (IN STRAIN PLY).    |
| FT | VARIANT  | 205 | 205   | S -> G (IN STRAIN PLY).    |
| FT | VARIANT  | 239 | 239   | I -> V (IN STRAIN PLY).    |
| FT | VARIANT  | 243 | 243   | S -> S (IN STRAIN PLY).    |
| FT | VARIANT  | 250 | 250   | S -> A (IN STRAIN PLY).    |
| FT | VARIANT  | 257 | 257   | N -> S (IN STRAIN PLY).    |
| FO | SEQUENCE | 325 | 35480 | MM; D1928A12AAB1C158 CRC64 |

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 8.0%;            | Score 184;         | DB 1;      | Length 325; |
| Best Local Similarity | 26.1%;           | Pred. No. 0.00012; |            |             |
| Matches 93;           | Conservative 56; | Mismatches 135;    | Indels 72; | Gaps 18;    |

|    |  |     |  |     |
|----|--|-----|--|-----|
| OY |  | 77  | GLNNAIYTAGETVNDNFSSNSMEAIQOALALNGGETTSVDST-----ALEGKYSILAN   | 132 |
| Dd |  | 13  | GIPVSAYIVNCTEPDCKSFNSALNGIKAKIEERIELVLYESSTNSYLSDLE---LKD    | 68  |
| OY |  | 133 | TNRWVWLSGFOHD-AFTRMUKIPEKNKOLETEKNIILIGIMTPTDENVIPTGRVINLT   | 191 |
| Dd |  | 69  | AGSNLMILIGRFSFDVAKAVALSLOPELKAYI-----IDPVSEEPITP-NLVGMF      | 118 |
| OY |  | 192 | KTEAGWLAGYAANSFLAKFPSPDPTRKSAIVYGCGISPAVTD-FLAGLAGIAKANLNK   | 250 |
| Dd |  | 119 | RAQGAEPLTGVIYAAKV-----SKTKGIIGLIGIEGEIVAFRYGYEACAKYAN--      | 166 |
| OY |  | 251 | SDKTKKITTTOKI-----EINLGFDVODSTKERLEQIASNKPKSTLLAVGPLTEIFSDII | 306 |
| Dd |  | 167 | --KQIKTSARIHSGFNDEVEG---RSVAITMYSDCI-----DIITHAASJGGICATV    | 214 |
| OY |  | 307 | ANO--NDRYLIGVDPSDLVYTTRKNKFPTSILNLGYSVSYLSDLYTRKKSNSRLACF    | 364 |
| Dd |  | 215 | AKEIGSGHYIIGVEDGESYL---APNNIITSARKDVGRSL-NIFTSNYLKTMT-----F  | 264 |
| OY |  | 365 | EFGKKSATVYIIGIKDFVPDIADTSLSCGNKKLATFEAISKAKEEKETVTTPAEVY     | 420 |
| Dd |  | 265 | EGGR---LINGYAKEGVYGFVK---NPRMP-----FELEKEIDMLSSKIINOEI       | 308 |

RESULT 3

|    |   |           |      |         |
|----|---|-----------|------|---------|
| ID | TMPC_TREPA  | STANDARD; | PRT; | 353 AA. |
| AC | P29724; O83339;   |           |      |         |
| DT | 01-APR-1993 (Rel. 25, Created)  |           |      |         |
| DT | 15-DEC-1998 (Rel. 37, Last annotation update)   |           |      |         |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update)   |           |      |         |
| DE | Membrane lipoprotein tmpc precursor (membrane protein C) (35 kDa antigen) (Lipoprotein Tpn35).  |           |      |         |
| GN | TMPC OR TP0319.   |           |      |         |
| OS | Treponema pallidum.   |           |      |         |
| OC | Bacteria; Spirochaetales; Spirochaetaceae; Treponema.   |           |      |         |
| OX | NCBI_TaxID:160;   |           |      |         |
| RN | [1]   |           |      |         |
| RP | SEQUENCE FROM N.A.  |           |      |         |
| RC | STRAIN-NICHOLS;   |           |      |         |
| RX | MEDLINE-91372962; PubMed-1894360;   |           |      |         |
| RT | Schouls L.M., van der Heide H.G.J., van Embden J.D.A.;  |           |      |         |
| RT | "Characterization of the 35-kilodalton Treponema pallidum subsp. pallidum recombinant lipoprotein tmpc and antibody response to lipidated and nonlipidated T. pallidum antigens."   |           |      |         |
| RL | Infect. Immun. 59:3536-3546(1991).  |           |      |         |
| RN | [2]   |           |      |         |
| RP | SEQUENCE FROM N.A.  |           |      |         |
| RC | STRAIN-NICHOLS;   |           |      |         |
| RX | MEDLINE-98332770; PubMed-9665876;   |           |      |         |
| RA | Frazer C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodson R., Winn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J., Khalak B., Richardson D., Howell J.K., Childabaram M., Uitterback T., McDonald L., Atliach P., Bowman C., Cotton M.D., Fujii C., Garland S., Hach B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C., |           |      |         |





```

RP SEQUENCE FROM N.A.
RC STRAIN-PBI;
RX MEDLINE-98010210; PubMed-9350727;
RA Roesler D., Hauser U., Wilske B.;
RT "Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnosis."
RL J. Clin. Microbiol. 35:2752-2758(1997).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).
CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL: X81518; CAA57238.1; -
CC InterPro: IPR003760; Bmp.
CC Pfam: PF02608; Bmp.1.
CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC Membrane; Lipoprotein; Signal.
CC KW SIGNAL 1 14 PROBABLE.
CC FT CHAIN 15 341 BASIC MEMBRANE PROTEIN B.
CC FT LIPID 15 15 N-ACYL DIGLYCERIDE (PROBABLE).
CC SQ SEQUENCE 341 AA; 37236 MW; 442BFBFBEDDC9A CRC64;

Query Match 7.4%; Score 170; DB 1; Length 341;
Best Local Similarity 26.8%; Pred. No. 0.00096;
Matches 91; Conservative 56; Mismatches 116; Indels 76; Gaps 20;

QY 87 GYVNDNSFNQSWEAIOQLALGGETSVDSSTAELEGYV---LANTNKN---VW 138
DB 36 GVDDDSFNSANRALRLREDEPENETKVFSSNA--SGVYSVSDLDLNMKNGSDLIW 93
QY 139 VLSGFQMGDAFTRMKLPENKOLFETENIILIG-IDMTDENYVPGRYINILPYKTEAG 197
DB 94 -LVGYMITDA-----SLVSLNPKISTYGIIDPYSDDOVQIPNLIIGVFRIDGA 143
QY 198 WLAGYANASFLAKKPSDPTRKSAIVIGGISPAVTD-FLAGYLAGIKANLNKNSDKRTK 256
DB 144 FLAGY-----IAKKSVS-----GKIFIGVKGDIYDAFRGYEAG-----AKYADGIE 189
QY 257 ITTDK-----IEINLGFVQDTSTKRELEQIASK--DKPSTLLAVAGPLTEFSDITANQN 310
DB 190 IVSEYSNSFSQDNING-----RAIANKKYAKAGIDIIHFMAGLAGI--GVIEAPK 235
QY 311 D-----RYLIGVDVDSLVYRTKTKKPFSTSLKNGYVSFVLSLDLYTKKNSRNLAGFEF 366
DB 236 ELGEGYIVTADODQSHL--APRNFTTSYIAKNVGDALYLYTSE-YLKNNT-----WEG 286
QY 367 GKKSATVYLGIKRFDYDIADTS---LEGNDKKLATAEAI 401
DB 287 GK---IIOGRLRDGVGLSNANKFEYIKVIERKIVNEEI 322

RESULT 7
BMPD_BORBU STANDARD; PRT; 341 AA.
AC 044743;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Basic membrane protein D precursor.
GN BMPD OR BB0385.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-JD1;
RX MEDLINE-96178617; PubMed-8606088;
RA Ramamoorthy R., Povinelli L., Philipp M.T.;
RT "Molecular characterization, genomic arrangement, and expression of bmpD, a new member of the bmp class of genes encoding membrane proteins of Borrelia burgdorferi."
RL Infect. Immun. 64:1259-1264(1996).
RN [2]
CC SEQUENCE FROM N.A.
CC STRAIN-ATCC 35210 / B31;
CC MEDLINE-98065943; PubMed-9403685;
CC Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
CC Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
CC Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
CC Peterson J., Karpavage A.R., Quackenbush J., Salzberg S., Hanson M.,
CC van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
CC Ullrichback T., Wathley L., McDonald L., Artlisch P., Bowman C.,
CC Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
CC Smith H.O., Venter J.C.;
CC "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."
CC Nature 380:580-586(1997).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).
CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL: U35450; AAC43984.1; -
CC DR EMBL: AE001144; BAB91505.1; -
CC DR TIGR: BB0385; -
CC InterPro: IPR003760; Bmp.
CC Pfam: PF02608; Bmp.1.
CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC Membrane; Lipoprotein; Signal; Complete proteome.
CC KW SIGNAL 1 16 PROBABLE.
CC FT CHAIN 17 341 BASIC MEMBRANE PROTEIN D.
CC FT LIPID 17 17 N-ACYL DIGLYCERIDE (PROBABLE).
CC SQ SEQUENCE 341 AA; 37163 MW; 63F6638F81A91D5 CRC64;

Query Match 7.4%; Score 169.5; DB 1; Length 341;
Best Local Similarity 25.4%; Pred. No. 0.001;
Matches 88; Conservative 64; Mismatches 137; Indels 57; Gaps 20;

QY 65 IMAKAAKANKHFGIMAVIYAGTVNDNSFNQSWEAIOQLALGGETSVDSSTAELE 124
DB 15 VACSSDDGKSEAKTVSLI-VDGAFDDKGFENSSSKAIRKKALDNLINIIIEKASTGNSYL 73
QY 125 KYSSILANTNKN-VWVLSGFQMGDAFTRMKLPENKOLFETENIILIGIMPTDENVIPT 183
DB 74 GDIANLEDGNSNLIMGI-GPRLSDIL--FQASASNVSV-----WAI--IEGVDEIDIPK 124
QY 184 GRYINLTYTEAGWLAGYANASFLAKKPSDPTRKSAIVIGGISPAVTD-FLAGYLAG 242
DB 125 -NLINISFRSEVAFVLAG-----FASKASKTGKIGFVGVRGVKLVESFMYGYEAG 174
QY 243 IKANNLNKSDKKTKITTDKI-----EINLGFVQDTSTKRELEQIASKDKSTLLAVAGPL 298
DB 175 AKYAN-----SNIKVVSQYVGTGFDGFG-----RSTASNN-----YRDGVYDIIFAAG-L 219
QY 299 TEIFSDIILANO--NDRYLIGVDVDSLVYRTKTKKPFSTSLKNGYVSFVLSLDLYTKK 356
DB 220 SGCIIVIAKAEKGDHIIIGVDQSYL--APRNIVYSAVKKYSLSL-----TKKY 271
QY 357 NSRNLAGFEFGKKSATVYLGIK-DREVDIADTSLEGNDKKLATAEAI 401

```

[illegible]

| DB         | 19   | ILGAGCSSSSDDKSSDDKSSKDF--TVAAVDTGTGVDHRSFNQSAAMEGLQKRGKANDMK  | 76                             |
|------------|--|---|--------------------------------|
| QY         | 110  | GGEITSVDSSTALEGKTKSSSLANTKKNVYLS-----GFQHGDAFTPRMLK-IPENKQ    | 160                            |
| Db         | 77   | GTGDQNYNYLQASAEADYK-----TYLNLNFAVRSDYDLIGCYKLDALIEEVSQKPRKNPF | 130                            |
| QY         | 161  | LFTEKNITIIIGIDMTDTEENVIPTRGRINLTYTEAGWLAGVANAASFLLAKRPSDPTKRS | 220                            |
| Db         | 131  | AIYVDPT-----DDRDNVVSIG-----FNDNDOSYLVGVAGI-----TTTKTN         | 168                            |
| QY         | 221  | AIIVGIGISPAVTD-FIAGYLLAGIKAMNLKNSDKRKTITTDKIEINLGFVQ---DTSRK  | 276                            |
| Db         | 169  | KVGFVGKVGKGYIDRBEAGFGTAGVKA-----VFNNAQIDVGYANDFAKA            | 212                            |
| QY         | 277  | ERLEDTASKDKRS---FLLAVAGPLTEFSPITIANONDR-----YLIGVDTDQ-----    | 321                            |
| Db         | 213  | DKGQDIASSMYSVGVDYIFHAAGTGNGVFAR-AKNLKKDKPSRAVWYIGVDRDQDECK    | 271                            |
| QY         | 322  | SLVYTKTKNKRFETSLKMLGYSVPSVLSDTLTKKSNRNLNG-FEEGKSKATVYLGIND    | 379                            |
| Db         | 272  | VTANGKDYNTWTLSEIKRVDAV---EDLKT-----NAKAGDPGPKG---IEYGLDK      | 319                            |
| QY         | 380  | RFVADIATSLKNDKRLKATEALISEAKKEEETKTIPEAEVKEETLEIEMP            | 430                            |
| Db         | 320  | DAVGISE-----HODNISNDVLAKE---EYKQKIVDGD-----IKVPEKP            | 357                            |
| RESULT     | 9  |   |                                |
| BMPA_BORAF |  |   |                                |
| ID         | BMPA_BORAF   | STANDARD: PRT: 339 AA.  |                                |
| IC         | 031280: 031281: 031282: 031283:  |   |                                |
| DT         | 15-DEC-1998 (Rel. 37, Created)   |   |                                |
| DT         | 15-DEC-1998 (Rel. 37, Last sequence update)  |   |                                |
| DT         | 15-DEC-1998 (Rel. 37, Last annotation update)  |   |                                |
| DE         | Basic membrane protein A precursor (Immunodominant antigen P39).   |   |                                |
| GN         | BMPA.  |   |                                |
| OS         | Borrelia afzelii.  |   |                                |
| CC         | Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.   |   |                                |
| OK         | NCBI_TaxID=29518:  |   |                                |
| RN         | [1]  |   |                                |
| RP         | SEQUENCE FROM N.A.   |   |                                |
| RC         | STRAIN=PKO, PMUDI, PLE, AND PLJ7:  |   |                                |
| RC         | MEDLINE=98010210: PubMed=9350727:  |   |                                |
| RA         | Roesler D., Hauser U., Wilske B.:  |   |                                |
| RT         | "Heterogeneity of BmpA (P39) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnosis.";  |   |                                |
| RL         | J. Clin. Microbiol. 35:2752-2758(1997).  |   |                                |
| CC         | -1- FUNCTION: NOT KNOWN: IMMUNOGENIC PROTEIN.  |   |                                |
| CC         | -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).   |   |                                |
| CC         | -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.   |   |                                |
| CC         | -----  |   |                                |
| CC         | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_licensese@isb-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_licensese@isb-sib.ch</a> ). |   |                                |
| CC         | -----  |   |                                |
| DR         | EMBL: X81516: CAA57236.1; -  |   |                                |
| DR         | EMBL: X97237: CAA65876.1; -  |   |                                |
| DR         | EMBL: X97243: CAA65878.1; -  |   |                                |
| DR         | EMBL: X97241: CAA65880.1; -  |   |                                |
| DR         | InterPro: IPR003760; Bmp.  |   |                                |
| DR         | Pfam: PF02608; Bmp. 1.   |   |                                |
| DR         | PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.   |   |                                |
| KW         | Antigen; Membrane; Lipoprotein; Signal.  |   |                                |
| FT         | SIGNAL   | 1   | 17                             |
| FT         | CHARIN   | 18  | 339                            |
| FT         | LIPID  | 18  | 18                             |
| FT         | VARIANT  | 125   | 125                            |
|            |  |   | BASIC MEMBRANE PROTEIN A.      |
|            |  |   | N-ACYL DIGLYCERIDE (PROBABLE). |
|            |  |   | A -> S (IN STRAIN PLE).        |



```

Db      286 YLKEGVGFV-----RNPKMISF-----ELEKIDNLSKKIKETIYP----- 325
Oy      435 DKQGESLOKLIND 447
       :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
Db      326 -SNKSEYEFKE 337

RESULT 11
ID      Y040_MYCGE          STANDARD:          PRT:          591 AA.
AC      P47286;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hypothetical lipidprotein MG040 precursor.
GN      MG040.
OS      Mycoplasma genitalium.
OC      Bacteria; Filumcutes; Bacillus/Clostridium group; Mollicutes;
        Mycoplasmataceae; Mycoplasma.
OX      NCBI_TaxID=2097;
[1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-ATCC 33530 / G-37;
RX      MEDLINE=96026346; PubMed=7569993;
RA      Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA      Fleischman R.D., Balt C.J., Kierlavage A.R., Sutton G., Kelley J.M.,
RA      Phillips J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA      Nguyen D.T., Utterback T.R., Saudak D.M., Phillips C.A., Merrick J.M.,
RA      Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA      Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT      "The minimal gene complement of Mycoplasma genitalium.";
RL      Science 270:397-403(1995).
RN      [2]
RP      SEQUENCE OF 448-517 FROM N.A.
RC      STRAIN-ATCC 33530 / G-37;
RX      MEDLINE=9407230; PubMed=8253680;
RA      Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT      "A survey of the Mycoplasma genitalium genome by using random
        sequencing.";
RL      J. Bacteriol. 175:7918-7930(1993).
-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
    (potential).
-1- SIMILARITY: SOME TO T.PALLIDIUM TMPC.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its contents is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; U39683; AAC71256.1;
EMBL; U02125; AAD12400.1;
TIGR; MG040;
InterPro: IPR003760; Bmp.
Pfam: PF02608; Bmp.1.
PROSITE; PS00013; PROKAR_LIPOPROTEIN_1.
Hypothetical protein: Lipoprotein; Membrane; signal;
Complete proteome.
SIGNAL              1           30
FT CHAIN            31           591     POTENTIAL.
FT LIPID             31           31     N-ACYL DIGLYCERIDE (POTENTIAL).
FT FT                31           31
SEQUENCE 591 AA; 64019 MW; 14A8FA31BB7E0928 CRC64;

Query Match      6.8%; Score 156.5; DB 1; Length 591;
Best Local Similarity 21.0%; Pred. No. 0.014;
Matches 100; Conservative 84; Mismatches 134; Indels 159; Gaps 27;

```

|            |   |                         |   |         |
|------------|---|-------------------------|---|---------|
| Dd         |   | 33                      | ATKIDSSVQLLVSNFSTLADKSFSQMSYEGEIRSFKKSGVDLPDASQLO-EGNGL     | 91      |
| Oy         |   | 127                     | -----XSILANTMKNWVYLSFGHGAFTRLMKIPENKOLFTEKNIIILGI           | 172     |
| Dd         |   | 92                      | WKRPGETSLDRIATEFNINIKNGSDVIYATGFMQ-----QESLAITSDDIRE---     | 138     |
| Oy         |   | 173                     | DWTDENYIPTG-----RYNLTYKTEAGMLAGYANNS                        | 206     |
| Dd         |   | 139                     | -QSDESLAKTGFIIVDGAIEKEFKRNGCVPOFKSTPTNMISVAARSDGCSFLTCVATAV | 197     |
| Oy         |   | 207                     | FL-----AKRPSDPTRSAIVIG--GIS-PAYTDFLAGLAGITKANM-----L        | 248     |
| Dd         |   | 198                     | YLNLNQEYFLDKSGMSTNSSNNMETVSGVIALPSTLSFLNGRLGIATYENEVIYKHL   | 257     |
| Oy         |   | 249                     | KNS-DKTKRTITDK-----IEING-----PDVOD-----NST                  | 275     |
| Dd         |   | 258                     | SDAQSSMAQVTTSKOTVLKOLOVANGERRIKTKIMSPQSGDETINIQDHQSGSFSPT   | 317     |
| Oy         |   | 276                     | KERLEQIASK--DK-PSTLLAVALGPLTE-IKSDIITANDRYLIGVDTDQSL---YYTK | 327     |
| Dd         |   | 318                     | EPRATTIANNLIDKGVNAIIPACHTQNLVVQIARRQAHTFAVICVSADSQELLDINDAP | 377     |
| Oy         |   | 328                     | TKNKF-----FTSLIKNLGIVSVFSLDYLRKKSRSRLAFPE-FGKKSATVYIG-      | 376     |
| Dd         |   | 378                     | NKDJKMKGNKKTIIPSSI-KALDVAVESISTLT-EKSSQN--GYOGGYNN----IGT   | 428     |
| Oy         |   | 377                     | IKDRFVADIAT-----SLECNCKKLAEATSEIAEKFEFEETKT                 | 413     |
| Dd         |   | 429                     | VKNNSVGVSFAGYEFLIDPVFWKNTSSMQAMSLSLKANMASSSDNKKISEVATK      | 485     |
| RESULT     |   | 12                      |   |         |
| BMPB_BORAF |   |                         |   |         |
| ID         | BMPB.BORAF  | STANDARD:               | PRT;  | 341 AA. |
| AC         | O31284;   |                         |   |         |
| Dt         | 15-DEC-1998 (Rel. 37,   | Created)                |   |         |
| Dt         | 15-DEC-1998 (Rel. 37,   | Last sequence update)   |   |         |
| Dt         | 15-DEC-1998 (Rel. 37,   | Last annotation update) |   |         |
| De         | Basic membrane protein B precursor.   |                         |   |         |
| Gn         | BMPB.   |                         |   |         |
| Os         | Borrelia afzelii.   |                         |   |         |
| Oc         | Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  |                         |   |         |
| Ox         | NCBI_TaxId=29518;   |                         |   |         |
| Rp         | [1]   |                         |   |         |
| Rn         | SEQUENCE FROM N.A.  |                         |   |         |
| Rc         | STRAIN=PKO;   |                         |   |         |
| Rx         | MEDLINE=98010210; Pubmed=9350727;   |                         |   |         |
| Ra         | Roesler D., Hauser U., Wilske B.;   |                         |   |         |
| Rt         | "Heterogeneity of Bmpa (P19) among European isolates of Borrelia burgdorferi sensu lato and influence of interspecies variability on serodiagnosis."  |                         |   |         |
| Rl         | J. Clin. Microbiol. 35:2752-2758(1997).   |                         |   |         |
| Rl         | -1-SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (Probable).   |                         |   |         |
| Cc         | -1-SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.   |                         |   |         |
| Cc         |   |                         |   |         |
| Cc         | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.jlb-slb.ch/announce/ or send an email to license@slb-slb.ch). |                         |   |         |
| Cc         |   |                         |   |         |
| Cc         | EMBL; X81519; CAAS7239.1; .   |                         |   |         |
| Dd         | InterPro: IPR003760; Bmp.   |                         |   |         |
| Dd         | Pfam: PF02608; Bmp. 1.  |                         |   |         |
| Dd         | PROSITE: PS00013; PROKAR.LIPOPROTEIN; 1.  |                         |   |         |
| Kw         | Membrane; Lipoprotein; Signal.  |                         |   |         |
| Ft         | SIGNAL  | 14                      | PROBABLE.   |         |
| Ft         | CHAIN   | 15                      | BASIC MEMBRANE PROTEIN B.                                   |         |
| Ft         | LIPID   | 15                      | N-ACYL DIGLYCERIDE (PROBABLE).                              |         |
| So         | SEQUENCE  | 341 AA;                 | 1EBEA3EE3054ACDB1 CRC64;                                    |         |



Query Match 6.7%; Score 154; DB 1; Length 341;  
 Best Local Similarity 25.4%; Pred. No. 0.0094;  
 Matches 90; Conservative 56; Mismatches 125; Indels 84; Gaps 19;

QY 87 GYVNDNSFNQSSWEALIOALGTGSEITSVDSSTAELEGKYS--LANTKN--VW 138  
 D 36 GYLDDKSFSSANEALLRLKKDEPENIEKVFSSA--VSGYSSVSVDLKKNGSGLIW 93  
 QY 139 VLSFGHGP-AFTRMKTIENKOLFTEKNII--ILGIDMTDENVIPTGTIVMLTKTEE 195  
 D 94 -LVGVLAVLSISVSLNEP-----INVGIDPIYGDVQIPVNL-----IGVERIEQ 141  
 QY 196 AGVLGAYANASFLAKKFPSPDPTKRSATVIGGSPAVTD-FINGYLAGIKAMN-----L 248  
 D 142 GAFLAGY-----IAKKSYS-----SKIGFLGKGDIVDAFRGYENGAKYANKGIEIYS 192  
 QY 249 KNSDKTKTTTDKIEIN---LGFVDVDTSTKRLQIASKDKPSTLLAVAGPLTEIFSD 304  
 D 193 EYSNSFSDIDIAVMANKYSKIGIDIHFPA--GLAGVGYEAPKEI-----237  
 QY 305 ILANQDRLLIGVDQDSLVYTKTKNKFSTILKNIGYSVFSVLSQLYTKKSNRLAGF 364  
 D 238 ---GGYIVVIGADQDQSHLAPK---NFTTSVIAKNVGDALYLITSE-----SLKDNVW 284  
 QY 365 ERGKKSATVYLGIDRFVDIADTSLGNDKLLATEAISEAKKEFEETKTIPAE 419  
 D 285 EGGK---IYOMGLRDGVGLSNAN-----EFEXIKDLERKINKKEIIVPCNQ 328

## RESULT 13

BMP\_C\_BORBU STANDARD; PRT; 353 AA.

AC 050169: 044859;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Basic membrane protein C precursor.  
 GN BMC OR B80384.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 ON NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=297;  
 RX MEDLINE=95080623; PubMed=7988902;  
 RA Aron L., Aleksun M., Perlee L., Schwartz I., Godfrey H.P.,  
 Cabello F.;  
 "Cloning and DNA sequence analysis of bmpC, a gene encoding a  
 potential membrane lipoprotein of Borrelia burgdorferi.";  
 FEMS Microbiol. Lett. 123:75-82(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=297;  
 RX MEDLINE=97132632; PubMed=8978084;  
 RA Aron L., Toth C., Godfrey H.P., Cabello F.C.;  
 "Identification and mapping of a chromosomal gene cluster of Borrelia  
 burgdorferi containing genes expressed in vivo.";  
 FEMS Microbiol. Lett. 145:309-314(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE=98065943; PubMed=9403685;  
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 Dougherty J., Tomb J.-F., Fleischmann R.D., Richardson D.,  
 Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,  
 van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
 Ullrichback T., Matthey L., McDonald L., Artlich P., Bowman C.,  
 Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 Smith H.O., Venter J.C.;  
 "Genomic sequence of a Lyme disease spirochete, Borrelia

RT burgdorferi.";  
 RL Nature 390:580-586(1997).  
 CC 1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
 CC (Probable).  
 CC 1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U49938; AAC44711.1; -;  
 DR EMBL: AF001143; AAC66756.1; -;  
 DR TIGR: B80384; -;  
 DR InterPro: IPR003760; Bmp.  
 DR Pfam: PF02608; Bmp; 1.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Membrane; Lipoprotein; Signal; Complete proteome.  
 FT SIGNAL 1 16 PROBABLE  
 FT CHAIN 17 353 BASIC MEMBRANE PROTEIN C.  
 FT LIPID 17 17 N-ACYL DIGLYCERIDE (PROBABLE).  
 FT VARIANT 179 179 V -> L (IN STRAIN 297).  
 SQ SEQUENCE 353 AA; 39824 MW; 2E8PF607D6CAB9B0 CRC64;

Query Match 6.4%; Score 147; DB 1; Length 353;  
 Best Local Similarity 23.7%; Pred. No. 0.027;  
 Matches 84; Conservative 53; Mismatches 131; Indels 86; Gaps 18;

QY 83 VTAGGYVNDNSFNQSSWEALIOALG-----ALRGELITSVDSSTAELEGK 127  
 D 34 VLAHSGYDDKGYNSVHDGVYKLRDNFGIKLTKSLRPYIEKRLUTDE--AMTEDAT 91  
 QY 128 SSLANTNKNVWVLSGPHGDAFTRWLKIPEKNOLFTEKNIIIGI---DWTDTENVIPT 183  
 D 92 EYQKNPLNFWL-LGYRFSQ-----SVLSTERPIYIGLIDAFYGDIG--VPR 139  
 QY 184 GRINLTYTEEGGWLAVGYNASFLAKKFPSPDPTKSAIYIG---GGISPAVDTFAGYL 240  
 D 140 NS-LAIKFRNEEAFVLAGYIAAKMSRKE-----KIGFLTGMSRSHVDFRFGK 187  
 QY 241 AGIKAMN--LKNSDKTKTTTDKIEINLGFVDVDTSTKRLQIA-----SKDKPSTLLAV 294  
 D 188 AGIFYANPKRLVSKKAPSLFD-----KEGKMAALFMKEDVGVYIFPI 232  
 QY 295 AGPLTEIFSDIIANQDR-VLIGVDVDTQSLVYTKTKNKFSTILKNIGYSVFSVLSQLYT 353  
 D 233 AGITGLGVYAAKELGPKYVIGLNDQSYI---APQNVITSLIKIGKAYISISEYI- 288  
 QY 354 KKSNSRLAGFEFGKKSATVYLGIDRFVDA-DTSLGNDKLLATEAISEAKK 406  
 D 289 ---NNRVFKG-----GIIDRGLKRGVIRIVADPVLNN--RLVDEVIDLENK 331

## RESULT 14

PEPB\_STRPY STANDARD; PRT; 498 AA.

AC 099XS1;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable dipeptidase B (EC 3.4.-.-).  
 GN PEPB OR SP22066.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillius/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 ON NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;

CC MEDLINE-21192684; PubMed-11296296;  
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
 RT "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
 CC -1- CATALYTIC ACTIVITY: DIPEPTIDASE + H(2)O = 2 AMINO ACID.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY U34.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AE006627; AAK34726.1; -  
 SQ Hydrolyase; Dipeptidase; Complete proteome.  
 SEQUENCE 498 AA; 55499 MW; B9394DC94C19AE4D CRC64;

Query Match 6.0%; Score 139; DB 1; Length 498;  
 Best Local Similarity 20.3%; Pred. No. 0.13;  
 Matches 110; Conservative 73; Mismatches 180; Indels 180; Gaps 26;  
 Oy 8 NKFELGLVPLSA--IATISAGW---DKETKE----- 37  
 Db 3 NKKISGLVSLITATSLASVSACGFIIGKDLTKDGSILGRTEDLEPHNKNKIVRLA 62  
 Oy 38 -----EKSDNQN-----KQTDVSKISGLVNERKEIMAAKANKHFGCL 78  
 Db 63 KDNPAEGWMDLSNGFEYPLPEHSYRSAPDPYTPNKGYYDE-----AGFNE-FCV 112  
 Oy 79 NMAIYTAGTVND-----NSFNOSWEAIOQLALGCE---ITSVDSSTALE 124  
 Db 113 SMS-ATVSASANDAIQKIDPIYKNGLAESSMTSVLLPSYKTAKEVALAKIVTEKGAEE 171  
 Oy 125 GYSSSLANTNKNW---VLSGFQHDGDAFTRWLIKIPENKOLFTEKNIILIGIDMTDENYI 181  
 Db 172 GNIVTLADKD-GIWMELISGHQYVA-----IKFPDCKAYVPPNFYLGHDVDFNKENKI 225  
 Oy 182 PGRINILTYTEEGAGWLAGYANASFLAKF--PSDPTKRSATVIGGIGSPAVTPIAGY 239  
 Db 226 ASEDEVKAKKRSYTEVDGKEH---IAKSYNPPLNDANRSKRSFSG----- 268  
 Oy 240 LAGIKAMWLKNSDKTKIT-----TDKIEINLGFDVQDSTKER----- 278  
 Db 269 -----IKSLDPDSKAVTYKDSNTELLQSTDKT-----FSLDDAMKLDNRNREFGLDKP 315  
 Oy 279 LEQIA--SKDRPSTLLAVAGPLTEIFSDIANQN--DRYLGVDYD-----OSLYYT 326  
 Db 316 LDQMLADGKGRKSKKAVKG---YAVPISPNVMEAHIFQLKIDIPALGCGWMLSLG 370  
 Oy 327 KTKNKFPSILKNLG--VSFVSVDLYTKSNSRNLAGFEGGKSAVYVIGIKRPFVDI 384  
 Db 371 SPRNAPYLPYGLNIRTEAYOEKSTOYNDKS-----WYTVSHINDL 413  
 Oy 385 ADTSLEGDKKLATEAISEAKKEPEKKTIPAEVKTLEIPMPDQOPDQOQSLDKL 444  
 Db 414 ----AAHPKPPGTAVIDEMK---GLEKWTIAEODKSTKISLVDVSPKAAQEKADKI 465  
 Oy 445 ITD 447  
 Db 466 SLID 468

RESULT 15  
 SCA4\_RICFE STANDARD: PRT: 981 AA.  
 ID SCA4\_RICFE  
 AC 09AJ37;  
 DT 16-Oct-2001 (Rel. 40, Created)

DT 16-Oct-2001 (Rel. 40, Last sequence update)  
 DT 16-Oct-2001 (Rel. 40, Last annotation update)  
 DE Antigenic heat-stable 120 kDa protein (PS120) (120 kDa antigen)  
 DE (Protein PS 120) (Fragment).  
 GN SCA4 OR D.  
 OS Rickettsia felis (Rickettsia azadi).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.  
 OX NCBI\_TaxID=42862;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sekeyova Z., Roux V., Raoult D.;  
 RT "Phylogenetic analysis of Rickettsia spp. by comparing the sequence of  
 RT gene D coding for an intracytoplasmic protein.";  
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AF16973; AAK31304.1; -  
 DR Antigen.  
 KW NON-TER  
 FT NON-TER  
 SQ SEQUENCE 981 AA; 107514 MW; 7F18F421E2C262EL CRC64;

Query Match 5.8%; Score 132.5; DB 1; Length 981;  
 Best Local Similarity 23.0%; Pred. No. 0.84;  
 Matches 108; Conservative 71; Mismatches 162; Indels 129; Gaps 26;  
 Oy 15 LVFPLSATAT-----ISAGCMDEKTTKEKSDNQNQKQTDVSK-ISGLVNERKSEI 65  
 Db 393 LLYATSAISLQMDLLNYYNAG---LTKER---DGNKQIDLINENATAILNNEKSDI 443  
 Oy 66 MAKADANKHFGELNMAIYTAGTVNDSFNOSWEAIOQLALGCEITVSVAELLEG 125  
 Db 444 --AEKQAN-----IALTEVTNNNDLTPPTKAA---GVNAVLETKIKDQNTPDLE- 489  
 Oy 126 KYSSSLANTNKNWVLSGFQHDGDAFTRWLIKIPENKOLFTEKNIILIGIDMTDENYI 182  
 Db 490 -KSKMLEATVAIALNSEMD-----PROKQMLEKAVDV-GLSLKDDASVTAID 537  
 Oy 183 --TGRIINILTYTEEGAGW---AGYANASFLAKFPPSDPTKRSATVIGGIG---SPA 231  
 Db 538 GITDAVIRKSNLSTEDKGTMLNAVGDKNVNASLSN---AEKOKLLGSVLRKGVGAQVLSPE 594  
 Oy 232 VDTFAGYIAGIKAMWLKNSDKTKIT--TDKIEINLGFDP-----VODTSKERLEQIA 283  
 Db 595 QOQLMQONLDKITTAQOTRNA---QITEVQGLANPAANTTAKTAIDQNTVTKVLDSPFK 650  
 Oy 284 SKDKPSTLLAVAGPLTEIFSDIANQNDRYLIGVDTQSLVYTKTKNKFPSILKNLGYS 343  
 Db 651 AEIKETLESI---TKVAASPLNGQDK-----ADIVAGMGEA 685  
 Oy 344 VFSVSLDLYTKSNSRNLAGFEGGKSAVYVIGIKRDRVDIADTSLSEGNOKKLAT---E 399  
 Db 686 I-----ASHKTAAPTEKISTISVEKGVASSTIDL-----EDKRLMTKGLVE 727  
 Oy 400 AISEAKKEFE---EKTITPAEVRKTEIPMPKOPDK---QOESLDK 443  
 Db 728 GIYEGKANPEITSEKTKAV-SKGIIDKSTAIPE--DKQALKDAANEAALDR 774

Search completed: July 18, 2002, 15:30:19  
 Job time: 66 sec



**THIS PAGE BLANK (USPTO)**



```

Db 3 RSSRNAASSDVK-----TKAIVTDTGVDKSFNOSANEGLQANGKEHNLKNGCTYF 58
Oy 118 SSTAELE-----GKYSSLANTKNWVLSGFQHGDAFTRLKIPENKOLFTEKNII 168
Db 59 OSTSEADYANNLQAAAGSYNLIIFGV-----GFALNNA-----VKAAAEHNDLVNY 104
Oy 169 ILGIDMTPTENYIPGRYINLTYKTEBEAGWLAGYANASFLAKFPSPDKRSAYIGGCI 228
Db 105 LIDDIYIKDQKN-----ASVTFADNESGYLAGVAAK-----TKTKQVGVGGI 149
Oy 229 -SPAVTDFIAGYLAGIKAMNLSKSKTKITTDKIEINLGFVDQDSTKERLEQJASKDK 287
Db 150 ESEVTSRFPAGKACV-----ASVDSIVQYDIAG-SFG-DAAGKTTIAAAGYAGAD- 201
Oy 288 PSTLLAVAGPL-TEIFSDIIANQND-----YLIGVDTQSL-VYT-----KTKNKFFT 334
Db 202 --IYVQVAGGTGAGVFAEAKSLNESPENEKYVIGVDRDQEAEGKTYTSKDKESFVLY 259
Oy 335 SLKMLGYSVFVSLDLYTKKSNRNLAGFEFGKKSATVYLGIKDFVDIADTSLGNDK 394
Db 260 SLTKQVGTIVKDI-----SNKAERGFPGQ---VIVSLKDKGVDLAVTNLSEEGK 308
Oy 395 KLAIEAISEAKKEFEKTKIP 416
Db 309 K-----AVEDAKAKILDGSKVP 326

```

```

RESULT 2
US-08-396-957A-5
: Sequence 5, Application US/08396957A
: Patent No. 5780041
: GENERAL INFORMATION:
: APPLICANT: SIMPSON, WARREN; SCHWAN, TOM G.
: TITLE OF INVENTION: ANTIGENIC PROTEINS AND
: TITLE OF INVENTION: GENES ENCODING SAME OF BORRELIA BURGDORFERI.
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
: STREET: 345 PARK AVENUE
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: USA
: ZIP: 10154
: COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY DISK
: COMPUTER: IBM PC COMPATIBLE
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORDPERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/396,957A
: FILING DATE: 01-MAR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/020,245
: FILING DATE: 19-FEB-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/664,731
: FILING DATE: 05-MAY-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/487,716
: FILING DATE: 05-MAY-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: DOROTHY R. AUTH
: REGISTRATION NUMBER: 36,434
: REFERENCE/DOCKET NUMBER: 2026-4018054
: TELEPHONE: (212) 758-4800
: TELEFAX: (212) 751-6849
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 341

```

```

: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: HYPOTHETICAL: No
: ORIGINAL SOURCE:
: ORGANISM: Borrelia burgdorferi
: STRAIN: Sh-2-82
: INDIVIDUAL ISOLATE:
: DEVELOPMENTAL STAGE:
: HAPLOTYPE:
: TISSUE TYPE:
: CELL TYPE:
: CELL LINE:
: ORGANELLE:
: FEATURE:
: NAME/KEY: p39a
: LOCATION:
: IDENTIFICATION METHOD:
: OTHER INFORMATION: p39a protein sequence
US-08-396-957A-5

```

```

Query Match 7.8%; Score 179; DB 1; Length 341;
Best Local Similarity 26.9%; Pred. No. 2.6e-08;
Matches 101; Conservative 59; Mismatches 123; Indels 92; Gaps 25;

```

```

Oy 72 ANKHFGILMAIYTAGTVDNSFNOSMEAIQGLALTGGETSYDSSSTAELEGKYS-- 129
Db 24 SSKKIKISMV---DGVLDKSFNSANBALRLKKDPENIEEVS--CAISGVSSIV 78
Oy 130 --LANTNKN---VWVLSGFQHGDAFTRLKIPENKOLFTEKNII--ILGIDMTPTENVI 181
Db 79 SLDMLKRRGSDLI--LVGYMLTD--SLVASENPKI--SYGIIDPIYGDVOYIPEMLI 133
Oy 182 PGRYINLTYKTEBEAGWLAGY--ANASFLAKFPSPDKRSAYIG--GGISPAVTD-FI 236
Db 134 A-----VFEREAGAFIAGYIAAKSFSGK-----IGFGGMGNIVDAFR 174
Oy 237 AGYLAKAMNLSNDKTKITTDKIEINLGFVDQDSTKERLEQJASK--DKPSTLLAV 294
Db 175 YCESGAKYAN-KDIEITISEYNSFSVDIG-----RTIASKMNSKGDIVYHF 221
Oy 295 AGPLTEI-FSDIIANQND-RYLIGVDTQDLSYTTKPKFTSLKMLGYSVFVSLDLY 352
Db 222 AAGLAGIVTEAKMIGGYYVIGADQDS--YLAPKN-FITSYVKNIGDALYLITGE-Y 277
Oy 353 YKKSNSRNLAGFEFGKKSATVYLGIKDFVDIADTSLGNDKKLATEAISEAKKEFEK 412
Db 278 IKNNV-----WEGK---VQVMGLRDVIGLPPAN-----EF-EYI 310
Oy 413 KTIPEEYKTLPIR 427
Db 311 KYLERKIYKKEITVP 325

```

```

RESULT 3
US-08-396-957A-4
: Sequence 4, Application US/08396957A
: Patent No. 5780041
: GENERAL INFORMATION:
: APPLICANT: SIMPSON, WARREN; SCHWAN, TOM G.
: TITLE OF INVENTION: ANTIGENIC PROTEINS AND
: TITLE OF INVENTION: GENES ENCODING SAME OF BORRELIA BURGDORFERI.
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
: STREET: 345 PARK AVENUE
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: USA
: ZIP: 10154
: COMPUTER READABLE FORM:

```

MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/396,957A  
FILING DATE: 01-MAR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/020,245  
FILING DATE: 19-FEB-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/664,731  
FILING DATE: 05-MAY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/487,716  
FILING DATE: 05-MAY-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4018054  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 339  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Borrelia burgdorferi  
STRAIN: Sh-2-82  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
FEATURE:  
NAME/KEY: p39  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: p39, protein  
OTHER INFORMATION: sequence.  
08-396-957A-4

Query Match 7.0%; Score 162; DB 1; Length 339;  
Best Local Similarity 22.8%; Pred. No. 9.2e-07;  
Matches 85; Conservative 69; Mismatches 135; Indels 84; Gaps 18;

08 87 GTVNDSPNOSSEWALIOQLATGGETTSVDST---AELEGKYSLATNKNVWVLG 142  
Db 37 GTDDKSPNENALNGYKAAVEEKIELVLEKSSNSTLSDEG---LKGAGSDLTMLTG 92  
QY 143 FQHG-D-AFTRWLIKIPENKOLFTEKNTIILGIDWTDENVIPTGRYINLTYKTEAGWLAG 201  
Db 93 YRSVDAAKVAALQNPDMKVAI-----IDPIYSNDPIR-ANLVGMTFRAGEAFLTG 142  
QY 202 YAAASLAKFPSPDPTKRSALIVGCGISPAVTP-PIAGYIAGIKAMULKMSDKTKTTTD 260  
Db 143 YIAAKL-----SKTGKIGFAGIGECIEIVAFRYGEAGAKYAN-----KDIKISTQ 188  
QY 261 KI-----EINLGFVQDPTSTKERLEQIASKDKPSTLLAVAGPLTEIFSDITANO--NDRYL 314  
Db 189 YIGSPADLEAGRSVATRMYSDEID-----IIHHAAGLGICIGIEVAKELGSGHYI 238

QY 315 IGVDPDQSLVYTKTKNKEFTSIILKILGYSVFSLVSLDYTKRKSRLNAGEFGKRSATVY 374  
Db 239 IGVEDQDAVL---APDVIATSTTKDVGRL-----NIFT--SNHKTNTFEGGK---LIN 285  
QY 375 LGIKRFDVADIADTSLBENDKRLATEALISEAKKEFEETKTIIPAEVKTLEIPMPKOP 424  
Db 286 YGLKEGVAGFV-----BNPKMISF-----ELEKEIDMLSKTKINKIEIIVP----- 325  
QY 435 DKQOESLDKLTLD 447  
Db 326 -SNKSEYKFLKE 337

RESULT 4  
US-08-313-412-1  
Sequence 1, Application US/08313412  
Patent No. 6248583  
GENERAL INFORMATION:  
APPLICANT: Aron Ph.D., Lieselotte  
APPLICANT: Cabello M.D., Felipe  
APPLICANT: Godfrey M.D., Henry P.  
APPLICANT: Schwartz Ph.D., Ira  
TITLE OF INVENTION: Chromosomally-Encoded Membrane  
TITLE OF INVENTION: Protein of Borrelia burgdorferi  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
STREET: P.O. Box 1051, Clinton Square  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/313,412  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 35553/1020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1304  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 353 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-313-412-1

Query Match 6.3%; Score 144; DB 4; Length 353;  
Best Local Similarity 23.4%; Pred. No. 4.4e-05;  
Matches 83; Conservative 54; Mismatches 131; Indels 86; Gaps 18;

08 83 VTAGGTVNDSPNOSSEWALIOQLG-----ALTGGEISVSSTAELGKY 127  
Db 34 VLAHGSFYDDKGYNSVDGVYKLDNNGIKLITKSLRPYIEGKRLTYDE--AMTDAY 91  
QY 128 SSLANTKNKNVWVLSGFOHGAFTRWLIKIPENKOLFTEKNTIILGIDWTDENVIPT 183  
Db 92 EVQKNPLNLFMLI-GYRFSDL-----SVKLSYERPDIIYGCIIIDAFYGDIG--VPR 139  
QY 184 GRYINLTYKTEAGWLAGYANASFLAKFPSPDPTKRSALIVG---GGISPAVVDFTIGYL 240  
Db 140 NS-LAIKFRNEEAFLAGYIAAKMSRKE-----KIGFLTGPMSHLKDKFKGFK 187





```

1 CURRENT APPLICATION NUMBER: US/09/308,375
2 CURRENT FILING DATE: 1999-05-14
3 EARLIER APPLICATION NUMBER: EP9719636.4
4 EARLIER FILING DATE: 1997-09-15
5 NUMBER OF SEQ ID NOS: 3
6 SOFTWARE: FastSeq for Windows Version 3.0.
7 SEQ ID NO 2
8 LENGTH: 2285
9 TYPE: PRT
10 ORGANISM: Bacillus subtilis
11 US-09-308-375-2

```

|                           |       |                 |             |             |
|---------------------------|-------|-----------------|-------------|-------------|
| Query Match               | 5.3%  | Score 123.5     | DB 4        | Length 2285 |
| Best Local Similarity     | 19.3% | Pred. No. 0.083 |             |             |
| Matches 105; Conservative | 84;   | Mismatches 188; | Indels 167; | Gaps 23     |

```

OY      15  LVPFLSLAATISAGCWMKRETTKEKESADNOKQITDVSKISGVNKEKSEIMAAKDANK 74
Db      452  LNFNIANDSISIA-----DKLNEVONNAVITLIDLANSIRKAGSTAFVEGLND 502
OY      75  HEGLMALIVIA--GGTVNDNSF-----NOSSEALQOLG--ALTGETTSDVS 119
Db      503  LIGITVIASTRESGNIVGNSILKTFIFARIGNOSSIKALEOIGISYKTAGGAKKASDL 562
OY      120  TALEEGRYSSIANTNK-----NWVLVSGF-----QHGDAFTFM 152
Db      563  ISEVAGKWDLDLSQKQNTSIVAGIYQLSRFAMNNNFISAOAAKTAANSIGSAMSEQ 622
OY      153  LKIPENKQL-----FTEKNIIIGIDMTDENVI---PTGRYINL-TYKTEAGW- 198
Db      623  QKVYADSLQARVNRKLNNEFTE--FAIMASDAFISDGLIEFTQAAGSLTMASTGVKISVGL 680
OY      199  ---LAGYANASFLAKKRPSPDKSAIVIG-----GJISPA----- 231
Db      681  PLLAAVSTHTLLSK--NRTTASSILITRAMGOETLATAGLEAMTAAVAASRYKT 738
OY      232  -----VTDFIAGYLAGIKAMNKL--NSDKKTKITTTDKIEINLGPVDSTKERLEQT 282
Db      739  ALRGLIATSTLVGGFAAL--GMALESLLISFAEAKKAKADDEEOSQQTNVEAITTNK----- 792
OY      283  ASKRPSTLLAVAGPLEIFESDIJIANONDYLLGVDPDQSLVYTKTKNKFETSLKNLGY 342
Db      793  DSTDK-----LIQYKELEQKVESLISLSDGEQELQYTOQLAQGFPLVAKGYDS 842
OY      343  SVFSVSLDYLTKKNSBNLAGFEFGKKSATAVYLGIKDRFYDIADTSLGENDKRLATEAIS 402
Db      843  QGNAIL-----KTNKELEKAIENTKE--YIALKKQ-----ETRD 874
OY      403  EAKKEFEETKTI--PAEYFKTLEIPEMDK-----QPDKQOESIDKL 444
Db      875  SAKTTFEASKEIKSKDELKQYKQIADYNDKGRPKWDLIADDDYKVAADKAKOSMLKA 934
OY      445  ITDI 448
Db      935  QSDI 938

RESULT      8
US-09-268-347-36
; Sequence 36, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 2411
; TYPE: PRT

```

ORGANISM: Haemophilus influenzae  
US-09-268-347-36

|                       |                  |                 |             |              |
|-----------------------|------------------|-----------------|-------------|--------------|
| Query Match           | 5.3%             | Score 122;      | DB 4;       | Length 2411; |
| Best Local Similarity | 20.8%;           | Pred. No. 0.1;  |             |              |
| Matches 108;          | Conservative 65; | Mismatches 220; | Indels 126; | Gaps 23      |

```

QY 12 GLGVLEFPLSATIYSACCMDEETEEKESADNOMKOI-----TDVSKIG-----LVN 59
Db 303 GNGLVTKAKVAIDAVNKKAGWRKTT-----TANGQGDPAATASGNVTFEESGDGTASVTK 358
QY 60 ERKESEIMAAKKADANKHFGNLM-----AIVTAGFTVND----- 91
Db 359 DTNGNGITVKKDAVYGGGLKKRDSOKKIIVADTTALTALTVGGKVAELAKEDDKKLVNAGDLV 418
QY 92 NSFQSSWEATOOGLGALTLGEIITSVDSSPALEBCKYSLSLANTKNNWVL---SGOFHG-- 146
Db 419 TALGNLSMKAKAAEADDTDDALBEI -SKDQEVKAGEIYTRKAGKNLKVDDGANFTYSLO 477
QY 147 DAFTRWLKIIP-----ENKOLTEKNIIIL-----GIDWTDENVIPDGRYINLYTK 192
Db 478 DALGLTSLITLGGTTNGGNGAKPATVINKDGLITLPAGNGGTTGTWTSIVTXDKGIKAGNKAI 537
QY 193 TEEAGWLAGVANAASEFLAKKPPSDPTKRSALYIGGSIAPATDTFAGVLAGIKAMNKNSD 252
Db 538 TNVASGLRAYDANF-----DVLNNSAIDDLNHNVEDA-----YKGLNLNENNNAN 562
QY 253 KRTKITTDKIEI-----NLGFDVO-DTSTKERLEQIASKDPSTLLAVAGPLTEIFSDII 306
Db 583 KQPLVDTSTATVGDRLKRLGWNVSTKNGTKEESNQKODE--VLFTGASAAATVSK-- 637
QY 307 ANONDRYLI-----GVDPDQSLVYTKTKNKKFFTSILK--NLGYSVFESVLSLTYT 355
Db 638 -SEMGKHTIYVAETRAOSGLEKDGDTIKIKVDNQMPTWNLVYGNNGTAV-----T 688
QY 354 KKSNSRMLAGEFEKKSATYVLIIGIKDR -FVDIADTSLSEGDNKLATPAEISEAKKEFEKT 412
Db 689 K-----GGEFVTKGTAT-----DADGKQVYVKQATANDADKKVAT--VKDVATAINSAA 735
QY 413 KTIPAEYVRKLTLEIPEMDPKOPDKQOSLKLITIDIKN 451
Db 736 TFKVTENLTITSIDEDNPTDNGSKDALAGDLTLTKKAGN 774

```

RESULT 9  
 US-08-923-992A-4  
 Sequence 4, Application US/08923992A  
 Patent No. 6280738  
 GENERAL INFORMATION:  
 APPLICANT: Tai, Joseph Y.  
 APPLICANT: Blake, Milan S.  
 TITLE OF INVENTION: No. 6280738-1GA Fc Binding Forms of the Group B  
 TITLE OF INVENTION: Streptococcal Beta Antigens  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 STREET: 1100 New York Avenue, N.W., Suite 600  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/923, 992A  
 FILING DATE: 05-SEP-1997  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/024,707

FILING DATE: 06-SEP-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Eamond, Robert W.  
 REGISTRATION NUMBER: 32,893  
 REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 371-2600  
 TELEFAX: (202) 371-2540  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1104 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-923-992A-4

Query Match 5.3%; Score 121.5; DB 4; Length 1104;  
 Best Local Similarity 18.2%; Pred. No. 0.032;  
 Matches 86; Conservative 77; Mismatches 207; Indels 103; Gaps 13;

```

OY 38 EKSADNONQITDVSKISGLVNERKSEIMAKADANKHFGGLNMAIVTAGTVNDNFNS 97
DB 164 DEELNIOKQVEMAMQAITNEDKSMKIEDIRKO-----AQAADKKEDA 211
OY 98 SWEAIOQALGTGETTVSDSTAE-LEGKYSLANTNKNVWVLSGFOHGDFTFRLKIP 156
DB 212 EVKVRRELGLKLFSSFKAGLDQOIQEHVKKETSSEENTQ-----KVD 252
OY 157 E-----NKQFTEKNITIIIDTDTENVITPTGRYINLYTKTERAGVLAYANASFLAKKF 212
DB 253 EHYANSLQIAKQSLLEELKATTEGATQVKNQFLENAOKLEIOLPIKETWKLX-RA 310
OY 213 PSDPTKRSALIVIGGISPATVDFIAGYLAGIKAMNLKNSDKTKITTDKIEINLGFVD 272
DB 311 MESLEQVKEKELKHNSEANLODLVAKSEIIVRECKLNQSKNLPBKLE-----EENH 365
OY 273 TSTKERLEQIASKDKPSTLLAAGPLTEIFSDIIANONDRYLI--GVTDQSLVYTKTN 330
DB 366 SKLKQVVEHFRKKFKTSEQVT---PKRVKRDIAANENNQOKTELVSPENITVYEGEDV 422
OY 331 KP-----FTSLKLNGLSVFVSVDLYTKKNSRLAGFEFG---KSAIV 373
DB 423 KETVAKSDSKTTLDFSLTLTKYNPVSVDRISTNTKNTDHNKIAETITIKNLKLNQSYV 482
OY 374 YGIDRFVADIATSL-----EGNDKKLATEAS 402
DB 483 TIKANDDSGNVEKFTTIVQKKEQVKTPEQKHSTEQVNPQEPKSNKNOLOELIK 542
OY 403 EAKKEFEKTKTTPAEVVRKLTLEIPMPDKDPK-----QOESLDKLITDINK 450
DB 543 SAQOELKLEKAI--KELMEQPEIPSNPEYGIQKISWESQKEPIQALATSPFK 593
  
```

RESULT 10

US-08-790-912-4

; Sequence 4, Application US/08790912

; Patent No. 5976542

; GENERAL INFORMATION:

; APPLICANT: Weiser, Jeffrey N.

; APPLICANT: Plaut, Andrew G.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESS: PANITCH SCHWARZE JACOBS & NADEL, P.C.

; STREET: 1601 Market Street, 36th Floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19103-2398

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/790,912  
 FILING DATE: 29-JAN-1997  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/026,939  
 FILING DATE: 23-SEP-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Leary, Kathryn  
 REGISTRATION NUMBER: 36,317  
 REFERENCE/DOCKET NUMBER: 7600-401  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (215) 567-2020  
 TELEFAX: (215) 567-2991  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1861 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-790-912-4

Query Match 5.2%; Score 119.5; DB 2; Length 1861;  
 Best Local Similarity 18.1%; Pred. No. 0.11;  
 Matches 95; Conservative 63; Mismatches 172; Indels 195; Gaps 21;

```

OY 16 VFPLSAIATISACMDKETEKEESADNONQITDVSKISGLVNERKSEIMAKADANKH 75
DB 1355 VEVISTISLSLGADYRVRNNNEHRAEALNK-----FVEDNAQGTARQD--H 1401
OY 76 EGLNMAIVTAGTVNDNSFNQSSWEAIOQALGTGETTVSDSTAELEGKYSLANTNK 135
DB 1402 YVYWRIL-----DEQREKLYRNIILVYDAYKFGDDTVDAKATVEAQ----- 1443
OY 136 NWVVLSCPHGDATFRL-----KIPENKO--LTFEKIIIIIGDWTDTENVITPTGRYI 187
DB 1444 -----FDSSNPAMKVFEGPVGNKVHNRKGAVATGDSVYMGYRMLDKDGAITYTH-- 1494
OY 188 NLTYKTEEPAGMLAGYANASFLAKKPSDPTKRSATVIGGISPATVDFIAGYLAGIKAWN 247
DB 1495 EMTHDSNDNIYLGIGYGRSGCLPEF-----FAKGL-----Q 1526
OY 248 LKNSDKRTITTDKIEINLGFVDITSTKERLEQIASKDKPSTLLAAGPLTEIFSDIIA 307
DB 1527 APDHPDADATITVNSI---LKYDKNDASEKSRLOVL---DPTKRQNMDDL---KNYVH 1575
OY 308 NOND-----RYLIGV-----DTQDSLVYTKKKEF-----TSILKLN----- 340
DB 1576 NMEVYIMLEYLGMSIYNRLSDVQKVALRKIEKNKYVADADGNDVYATNVIKNTMDA 1635
OY 341 -----GYSEFVSVDLYTKKNSRLNAGFEFGKKSAT 372
DB 1636 OKLNSFNSLIENDLSAREYKNGDYERNCYHTIKLFSPIYSLASSEKGPGLMKRRILAY 1695
OY 373 VYL---GIDRFVADIAT-----SLEGNKKLATEAI-----SE 403
DB 1696 ELLAAKGFQDGVVPIYSNQVEDDAKONGKTTISYKTRGLVTDVLVLRKVFNGCFNNWTE 1755
OY 404 AKKE-FEEK-----TKTTPAEVVKTL 424
DB 1756 FRKMYEERKKNFDSLNVKTFDDTRQPTWTSYATKTIISTVEELQTL 1800
  
```

RESULT 11

US-09-336-447A-5

; Sequence 5, Application US/09336447A

; Patent No. 6310190

```

; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 892
; TYPE: PR1
; ORGANISM: Moraxella catarrhalis
US-09-336-447A-5
```

Query Match 5.2%; Score 119; DB 4; Length 892;

Best Local Similarity 20.8%; Pred. No. 0.039; Matches 105; Conservative 74; Mismatches 196; Indels 130; Gaps 23;

```

QY 23 ATTSACQWKEETTKERKSADNOKQIT-DVSKISGLVNERKSEIMAKKADANKHFGIMMA 81
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 146 STIGGGYNNQATGEKSTVAGRRNNQATGNNSTVAG-----GSYNQATGNNST 192
QY 82 IYTAGGTVDNSFNQSSWEAIQOLGALTGEIITSVDSSTAELEGKYSGLANTKKNVYLS 141
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 193 V--AGG-----SHNQATGE-----GSPAAGVENKANNANNAVALGKNNTIDGNS----- 234
QY 142 GFOHGDATFTWLKIPENKQLETER-NIIIGIDMTDEN-----VIP 182
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 235 -----VAIGSNNTIDSGKNVFIIGSS-TTNTNQSGLVGHNTAGKAKTAVS 262
QY 183 TGRYINLTYTEBAGWLACYANASFLAK--KPPSDPTKSAIYIG-GGISPAVTDPIAG 238
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 283 SAKVNGILT-----LGNFAGASKTGNGTVSVSENNENQIYVNGAGNISADSTDVNG 334
QY 239 -----YLAGIKAWNLKN-----SDKTKTKITDK--TEINL--- 266
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 335 SOLXALATAVKADADENFKALTYTQNTLILGEGAQDALIAQONOTDIYANKTAIERNPNRT 394
QY 267 ---GPDVQDTS---TERLE-QIASKDKPSTLLAVAGPLTEIFSDITANONDRYLIGVDT 319
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 395 VVNGFIEKKKAGIAKNQADIQLENNVGEELNLISGRLLDQKADIDNNINNIYDLAQOQ 454
QY 320 DQSLVYTKRKNKFTSLKWLGVSVF---SVLSDLTYTKSNSR---NLAGEFEGKKS- 370
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 455 DQSSSDIKTLKKNVEGLDLISGRLLDQKADLTKDKITLNNVVEEGLDLSGRLLDQKAD 514
QY 371 -ATVYIGIKDRFYDIAD---TSLEGNDKRLATEAISEAKFEFEKTKTIPAE--VRKT 423
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 515 IAKNQADIAQONOTDIQDLAAYNELQDQYAKQTEALDALNKKASANTDRATLAEELGIAEN 574
QY 424 LEIPEMPKQPDQKQESLKLITDI 448
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 575 KKDQAIKAQANENKDGIAKNQADI 599
```

```

RESULT 12
US-09-074-658-75
; Sequence 75, Application US/09074658
; Patent No. 6184371
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Run-Pan Du
; APPLICANT: Quljun Wang
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: LACTOPERRIN RECEPTOR GENES OF MORAXELLA
; NUMBER OF SEQUENCES: 78
```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 08-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1153
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 941 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-658-75
```

Query Match 5.1%; Score 117.5; DB 4; Length 941;

Best Local Similarity 19.5%; Pred. No. 0.058; Matches 79; Conservative 68; Mismatches 157; Indels 101; Gaps 17;

```

QY 32 KETTKERKSADNOKQITVDYSKISG-----LVNERK-----SEIMAKAD 71
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 256 KKKKAEYFLAEGSEELKPAKLAGNGLNQLNRVVEERKKNQSLAEEEMAVREAO 315
QY 72 ANKHFGILMAIYVAGGVYVNSFNQSSWEAIQOLGALTGE--ITSVDSSTAELEGYSS 129
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 316 A-RHENLSAQAIVTGGGILDPMDYRSGSWLAKIGYRGGHYVGGVFEPTQ--RYDT 371
QY 130 LANTKKNVWL-----GFOHGDATFTWLK-IPENKQLETKNII-LGIDMTDE 178
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 372 RDMTEKQYIGTDEKTRKSDSGVYDGDPRDGLYFVNIEEMGDKNLVVGIGIKYRTK 431
QY 179 NVIP--TGRYINLTYTEBAGWLACYANASFLAKKPPSDPTKSAIYIGGISPAVTDPI 236
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 432 FIDEHRRRRRGGILYRENEAYSPDNWADKAVLSFDKQGVATDNTLTLCNCAVYPSVDK- 489
QY 237 AGYLAGIKAWNLKNSDK-----KTKITDKIENLGGPDVQDTSKER 278
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 490 ACRASADKPYSDSDFFHYRBOHNVLNALLEKSLKKNKTKHHTLTGFGVD----- 540
QY 279 LEOYIAK--DKPSTLLAVAGPLTEIFSDITANONDRYLIGVDTQDQSLVYTKRKNKFTSI 336
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 541 ----ASKAVSRPEQLSHMARISE-FSDYADDGKYYKLLK-----PEV 579
QY 337 LKNLGVSVFSLDLYTKKSNRLAGFERGKKSATVYLCIKDRF 381
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 580 VEG---SVCGYIETLRSRKCVPRKING-----SNHISLDRF 614
```

```

RESULT 13
US-08-409-995-4
; Sequence 4, Application US/08409995
; Patent No. 5646259
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen I.
; APPLICANT: St. Gene III, Joseph W.
; TITLE OF INVENTION: Haemophilus Adhesion Proteins
```

```

: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
: STREET: Four Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/409,995
: FILING DATE: 24-MAR-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Silva, Robin M.
: REGISTRATION NUMBER: 38,304
: REFERENCE/DOCKET NUMBER: A-61053/RFT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 781-1989
: TELEFAX: (415) 398-3249
: TELE: 910 277299
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1912 amino acids
: TYPE: amino acid
: STRANDEDNESS: double
: TOPOLOGY: unknown
: US-08-409-995-4

```

Query Match 5.1%; Score 117; DB 1; Length 1912;

Best Local Similarity 21.0%; Pred. No. 0.2; Indels 128; Gaps 24;

Matches 109; Conservative 65; Mismatches 217;

```

OY 12 GIGLVEPLSAITTSAGCWDKETEKEKSADNQNQI-----TDVSKISG-----LVN 59
DB 306 GNGLTATAKAVIDAVNKAQKRVKTT-----TANGQNGDFATVAGSTVTEPSGGTTASVTK 361
OY 60 EKKSSEIMAKADANKHFGILN-----AIVTAGTVND----- 91
DB 362 DINGNGITVYKDAKVGDDIKFSDSKRIYADTTALTVTGKVAEIKEDDKKLVNAGDLY 421
OY 92 NSFNOSSWEAIDQALNGEITTSVDSSTAELGKYSSSLANTNKNVWYL---SGFOHG-- 146
DB 422 TALGNLSMKA--KABADIDGALGEGI--SKDOEVKAGETVTFKAGKMLKVKODGANFTYSIQ 478
OY 147 DAFTRMLKIP-----ENKQLFTEKNIIIL-----GIDMTDENVIPTGRYINLTYK 192
DB 479 DALTGLTSTLTGCTTNGGNDATVINKDGLTTTPAGNGCTGTNTISVTKDGIRAGNKAI 538
OY 193 TEEAGMLAGYANASFLAKKFPSPDKRSALIVGGISPAVTVDFIAGYLAGIKAMNLSKD 252
DB 539 TNVAGSLRAYDANF-----DVLNNSATDILNRHVEDA-----YKGLNLNNEKNAN 583
OY 253 KTKITTTDKIEI-----NLGPDVQ--DTSTKERLEQIASKDKPSTLLAVAGPLTEIFSDII 306
DB 584 KQPLVTDSTAAATVGDRLKIGWVSTKNGTKESSNOYKQADE--VLEFGAGAAVTYSK--- 638
OY 307 ANQNDRLYL-----GVDTDQSLVYTKTKNKFETSLIK--NLGYSVPSVLSLDTY 353
DB 639 -SENCKHHTTVSAETKADCGLEKDGDTIKLKVDNQNTDNVLTGNNNGAV-----T 689
OY 354 KKSNSRNLAGFEFGKKSATVYIGIKDR--FVIDIADTSLSEGNOKKLTAISEKKKEFEET 412
DB 690 K-----GGFETVKTGAT-----DADRGAVPVKDATANDADKAVAT--VKDVATAINSAA 736
OY 413 KTIPEAVRKTLIEIPMPDKOPDKQOESLKLITDINK 451
DB 737 TIVKTEINTLTSIDENPTDNGKDALAKAGDITLTKAGKN 775

```

RESULT 14  
US-08-685-467-4

; Sequence 4, Application US/08685467

; Patent No. 6060059

; GENERAL INFORMATION:

; APPLICANT: St. Gene III, Joseph W.

; APPLICANT: Barenkamp, Stephen J.

; TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: United States

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/685,467

; FILING DATE: 22-JUL-1996

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/409,995

; FILING DATE: 24-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Silva, Robin M.

; REGISTRATION NUMBER: 38,304

; REFERENCE/DOCKET NUMBER: A-61053-2/RFT/RMS

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; TELE: 910 277299

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1912 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; US-08-685-467-4

Query Match 5.1%; Score 117; DB 3; Length 1912;

Best Local Similarity 21.0%; Pred. No. 0.2;

Matches 109; Conservative 65; Mismatches 217; Indels 128; Gaps 24;

```

OY 12 GIGLVEPLSAITTSAGCWDKETEKEKSADNQNQI-----TDVSKISG-----LVN 59
DB 306 GNGLTATAKAVIDAVNKAQKRVKTT-----TANGQNGDFATVAGSTVTEPSGGTTASVTK 361
OY 60 EKKSSEIMAKADANKHFGILN-----AIVTAGTVND----- 91
DB 362 DINGNGITVYKDAKVGDDIKFSDSKRIYADTTALTVTGKVAEIKEDDKKLVNAGDLY 421
OY 92 NSFNOSSWEAIDQALNGEITTSVDSSTAELGKYSSSLANTNKNVWYL---SGFOHG-- 146
DB 422 TALGNLSMKA--KABADIDGALGEGI--SKDOEVKAGETVTFKAGKMLKVKODGANFTYSIQ 478
OY 147 DAFTRMLKIP-----ENKQLFTEKNIIIL-----GIDMTDENVIPTGRYINLTYK 192
DB 479 DALTGLTSTLTGCTTNGGNDATVINKDGLTTTPAGNGCTGTNTISVTKDGIRAGNKAI 538
OY 193 TEEAGMLAGYANASFLAKKFPSPDKRSALIVGGISPAVTVDFIAGYLAGIKAMNLSKD 252
DB 539 TNVAGSLRAYDANF-----DVLNNSATDILNRHVEDA-----YKGLNLNNEKNAN 583
OY 253 KTKITTTDKIEI-----NLGPDVQ--DTSTKERLEQIASKDKPSTLLAVAGPLTEIFSDII 306

```

```
Db 584 KQPLVYDSTAATVGDJRLKLGWVSTKNGTKEESNOVQADE--VLEFGAGAAVTYSK--- 638
QY 307 ANONDRILI-----GVDTDOSLYVTKTKNKFSTILK--NLGYSVFVLSLDLYT 353
Db 639 -SENGKHTITVSAEYKADCGLEKDDGPTIKLKYDNONTDNLVGNIGTAV-----T 689
QY 354 KKSNSRNLAGEFEGKKSATVYLGIKDR-FVADIADTSLEGNDKRLATPAISEAKKEFEKT 412
Db 690 K-----GGFEYVTKGAT----DADRGKVTYKDATANDADKKVAT--VKDVATAINSAA 736
QY 413 KITPAEEVKTLEIPEKPDKQPKQOESLDKLTITDINKN 451
Db 737 TFWKTENLTTSIDENPTDNGKDALKAGDTLTFKAGKN 775
```

## RESULT 15

```
US-09-377-155-33
; Sequence 33, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 2353
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; US-09-377-155-33
```

Query Match 5.1%; Score 117; DB 4; Length 2353;

Best Local Similarity 21.0%; Pred. No. 0.28;

Matches 109; Conservative 65; Mismatches 217; Indels 128; Gaps 24;

```
QY 12 GLGLVPLSLAATISAGCWMKETTKEESADNONKQI-----TDVSKISG-----LVN 59
Db 307 GNGLVTAKAVIDAVNKAGWVYKTT---TANGONGDPATVAGSTVTFESGDGTTASVTK 362
QY 60 ERKSEIIMAAKADANKHEGLNM-----AIVTAGTAVND----- 91
Db 363 DTNGNGITVYDAKVGGLGKLFDSODKIVADPTALTLYTGKVAELAKEDDKKLVNAGDLV 422
QY 92 NSFNOSWEAIIQOLGALTGCEITSVDSSTAELEGKYSILANTKNKVVYL--SGFOHG-- 146
Db 423 TALGNLSWKA--KAADPTDGLALGCI--SKDEYVAGETVTFKAGKNLKKODGANFTYSLO 479
QY 147 DATFRMLKIP-----ENKOLFTEKNIIIL-----GIDWTDENVIPTGRINILYK 192
Db 480 DALGLGLSTYLGTTNGGNDKATVINKDGLTTPAGNGGTTGTNTISVTKDGIRAKGNKAI 539
QY 193 TEEAGWLAGYANASFLAKKFPSPDPTKRSALVIGGSIAPATDFIAGYLAGIKAMNLKNSD 252
Db 540 TNVASGLRAYDANF-----DYLNNSATDILNRHVEDA-----YKGLNLNNEKNAN 584
QY 253 KTKKITTDKLEI-----NLGFDVQ--DTSTKERLEQIASKDKPSTLLAVAGPLTEIFSDII 306
Db 585 KQPLVYDSTAATVGDJRLKLGWVSTKNGTKEESNOVQADE--VLEFGAGAAVTYSK--- 639
QY 307 ANONDRILI-----GVDTDOSLYVTKTKNKFSTILK--NLGYSVFVLSLDLYT 353
Db 640 -SENGKHTITVSAEYKADCGLEKDDGPTIKLKYDNONTDNLVGNIGTAV-----T 690
```

```
QY 354 KKSNSRNLAGEFEGKKSATVYLGIKDR-FVADIADTSLEGNDKRLATPAISEAKKEFEKT 412
Db 691 K-----GGFEYVTKGAT----DADRGKVTYKDATANDADKKVAT--VKDVATAINSAA 737
QY 413 KITPAEEVKTLEIPEKPDKQPKQOESLDKLTITDINKN 451
Db 738 TFWKTENLTTSIDENPTDNGKDALKAGDTLTFKAGKN 776
```

Search completed: July 18, 2002, 09:14:54  
Job time: 34 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2002, 15:29:13 ; Search time 50.25 seconds  
(without alignments)  
1552.651 Million cell updates/sec

Title: US-09-676-249A-2  
Perfect score: 2299  
Sequence: 1 MKRKIKMKFLGLVPEPLS.....KQPDKQDESLDKLITDKKN 451

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Minimum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertibrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID     | Description        |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1          | 706   | 30.7        | 457    | 2  | Q48902 | Q48902 mycoplasma  |
| 2          | 554.5 | 24.1        | 461    | 16 | Q980L5 | Q980L5 mycoplasma  |
| 3          | 417   | 18.1        | 465    | 2  | Q9X775 | Q9X775 mycoplasma  |
| 4          | 400.5 | 17.4        | 428    | 2  | Q52311 | Q52311 mycoplasma  |
| 5          | 397.5 | 17.3        | 428    | 2  | Q9RGX5 | Q9RGX5 mycoplasma  |
| 6          | 394.5 | 17.2        | 428    | 2  | Q32417 | Q32417 mycoplasma  |
| 7          | 394   | 17.1        | 429    | 2  | Q9RGX6 | Q9RGX6 mycoplasma  |
| 8          | 394   | 17.1        | 429    | 2  | Q9RGX4 | Q9RGX4 mycoplasma  |
| 9          | 393.5 | 17.1        | 428    | 2  | Q9RGX7 | Q9RGX7 mycoplasma  |
| 10         | 389.5 | 16.9        | 428    | 2  | Q9RGX3 | Q9RGX3 mycoplasma  |
| 11         | 387.5 | 16.9        | 428    | 2  | Q9R3N6 | Q9R3N6 mycoplasma  |
| 12         | 332.5 | 10.1        | 350    | 16 | Q9R2H4 | Q9R2H4 streptococc |
| 13         | 229.5 | 10.0        | 349    | 2  | Q9E9V1 | Q9E9V1 clostridium |
| 14         | 223   | 9.7         | 350    | 16 | Q97RH0 | Q97RH0 streptococc |
| 15         | 207   | 9.0         | 350    | 16 | Q9CFM9 | Q9CFM9 lactococcus |
| 16         | 195.5 | 8.5         | 357    | 16 | Q97L60 | Q97L60 clostridium |

|    |       |     |      |    |        |                    |
|----|-------|-----|------|----|--------|--------------------|
| 17 | 187.5 | 8.2 | 516  | 16 | Q9PRD3 | Q9PRD3 ureaplasma  |
| 18 | 183   | 8.0 | 342  | 16 | Q9A1P7 | Q9A1P7 streptococc |
| 19 | 183   | 8.0 | 348  | 2  | Q9AK41 | Q9AK41 streptococc |
| 20 | 181.5 | 7.9 | 359  | 16 | Q9MXV7 | Q9MXV7 thermotoga  |
| 21 | 174   | 7.6 | 357  | 16 | Q92BW7 | Q92BW7 listeria in |
| 22 | 172   | 7.5 | 341  | 2  | Q9EPD3 | Q9EPD3 borrelia bu |
| 23 | 171   | 7.4 | 264  | 2  | Q31358 | Q31358 borrelia ga |
| 24 | 168.5 | 7.3 | 525  | 16 | Q9POR5 | Q9POR5 ureaplasma  |
| 25 | 165   | 7.2 | 347  | 2  | Q9AK42 | Q9AK42 streptococc |
| 26 | 164   | 7.1 | 524  | 16 | Q9PRC9 | Q9PRC9 ureaplasma  |
| 27 | 162   | 7.0 | 339  | 2  | Q9S6C1 | Q9S6C1 borrelia bu |
| 28 | 162   | 7.0 | 339  | 2  | Q87960 | Q87960 borrelia bu |
| 29 | 161   | 7.0 | 335  | 2  | Q9R776 | Q9R776 borrelia bu |
| 30 | 158   | 6.9 | 339  | 2  | Q9EPD2 | Q9EPD2 borrelia bu |
| 31 | 157   | 6.8 | 325  | 2  | Q9R777 | Q9R777 borrelia bu |
| 32 | 149   | 6.5 | 329  | 16 | Q9RGW4 | Q9RGW4 rhizobium l |
| 33 | 146   | 6.4 | 264  | 2  | Q31359 | Q31359 borrelia ga |
| 34 | 144.5 | 6.3 | 243  | 2  | Q31361 | Q31361 borrelia ga |
| 35 | 144   | 6.3 | 333  | 2  | Q93V09 | Q93V09 borrelia bu |
| 36 | 143.5 | 6.2 | 384  | 2  | Q46239 | Q46239 clostridium |
| 37 | 141.5 | 6.2 | 539  | 16 | Q9P010 | Q9P010 ureaplasma  |
| 38 | 141.5 | 6.2 | 889  | 2  | Q91961 | Q91961 moraxella c |
| 39 | 141   | 6.1 | 379  | 16 | Q9RSQ4 | Q9RSQ4 deinococcus |
| 40 | 136.5 | 5.9 | 912  | 2  | Q9XD52 | Q9XD52 moraxella c |
| 41 | 133.5 | 5.8 | 330  | 16 | Q926D6 | Q926D6 rhizobium m |
| 42 | 131.5 | 5.7 | 1786 | 5  | Q9U0P0 | Q9U0P0 plasmodium  |
| 43 | 130   | 5.7 | 484  | 16 | Q98RB2 | Q98RB2 mycoplasma  |
| 44 | 130   | 5.7 | 894  | 2  | Q91962 | Q91962 moraxella c |
| 45 | 130   | 5.7 | 1170 | 2  | Q9AJR8 | Q9AJR8 erysipeloct |

## ALIGNMENTS

| RESULT | ID   | PRELIMINARY | PRT | 457 AA. |
|--------|--|-------------|-----|---------|
| Q48902 | Q48902   | Q48902      |     |         |
| AC     | Q48902   | Q48902      |     |         |
| DT     | 01-NOV-1996 (TREMBLrel. 01, Created)                                 |             |     |         |
| DT     | 01-NOV-1996 (TREMBLrel. 01, Last sequence update)                    |             |     |         |
| DT     | 01-DEC-2001 (TREMBLrel. 19, Last annotation update)                  |             |     |         |
| DE     | AG 243-5 PROTEIN PRECURSOR.  |             |     |         |
| OS     | Mycoplasma arginalis.  |             |     |         |
| OC     | Bacteria: Firmicutes; Bacillus/Clostridium group; Mollicutes;        |             |     |         |
| OC     | Mycoplasmataceae; Mycoplasma.  |             |     |         |
| OX     | NCBI_TaxID=2094;   |             |     |         |
| RN     | [1]  |             |     |         |
| RP     | SEQUENCE FROM N.A.   |             |     |         |
| RX     | MEDLINE=96163149; PubMed=8551970;                                    |             |     |         |
| RA     | Ushio S., Iwaki K., Tanai M., Ohta T., Fukuda S., Sugimura K.,       |             |     |         |
| RA     | Kurimoto M.;   |             |     |         |
| RT     | "Metastatic promoting activity of a novel molecule, Ag 243-5 derived |             |     |         |
| RT     | from Mycoplasma and the determination of the complete nucleotide     |             |     |         |
| RT     | sequence."   |             |     |         |
| RL     | Microbiol. Immunol. 39:393-400(1995).                                |             |     |         |
| DR     | EMBL: D16674; BAA04082.1; -  |             |     |         |
| DR     | InterPro: IPR003760; Bmp.  |             |     |         |
| DR     | Pfam: PF02608; Bmp; 1.   |             |     |         |
| KW     | Signal.  |             |     |         |
| FT     | SIGNAL   |             |     |         |
| FT     | CHAIN  |             |     |         |
| FT     | SEQUENCE   |             |     |         |

Query Match 30.7%; Score 706; DB 2; Length 457;  
Best Local Similarity 38.4%; Pred. No. 8.7e-33;  
Matches 173; Conservative 87; Mismatches 145; Indels 46; Gaps 16;

|    |    |  |
|----|----|--|
| QY | 14 | GLVPLSAIAIT-ISAGCWDKETTKEESADNQNKQIDVSKISGLVNERSEIIMAAKADA 72    |
| DB | 12 | GTISTVASVATFVSCG---ETDKEGKII-----RIFD-----NSFVDRQAEIKAKA----- 54 |
| QY | 73 | NKHGELMAIVTAGTVDNSFNQSSWEAI---QQLGALTG-GEITSVDSSTALEGKY 127      |

```

Db 55 -NFPEFTVLLTAGTGVOQKSFNQSISWEAVLEHYDQIEKTTMLDRVSOETNNQSELIGKY 113
Qy 128 SSLANTNKNWVVLSCFOHDAFTRWLIKIPEN-----KQFETKNNIIIGIDW---TDTEN 179
Db 114 KRFPLNKNRWVLLITGGQOQOEFPKFKQTDNSGKKSDDLAEKVTIIVAVNDMLSEKEDD 173
Qy 180 VIPTGRYINLYKTEEAGMLAGYANASFLAKKPPSDPTRRSALVIGGISPATVDFIACY 239
Db 174 LKAGHFISLYKTEEBAGFIAGYASSKFLAYKFPNDEAKRTIAPFCGGAGAGATDFEIAEF 233
Qy 240 LAGIATNKNKNSDKTKITTDKIEIINLGFVDVDTSTKELEQIASDKRSTLLAVAGPT 293
Db 234 LAGIAKYNNDNPJAKVTIISDNNINIDTGF-ISNDKTATFINGIVKNS--SLVLPVAGSLT 290
Qy 300 EITFSDI--ANONDRYLIGVTDQSLVYTKTKNKPFTSTILKNLGYVSFVSLDYTKKSN 357
Db 291 SSVVDALKSNKNDTKLIVDQDSKIFSPA-TVFTSTIEKHIGRTIYVLTIDMLKKED 349
Qy 358 SBNLAGFEFGK---KSATVYLGIKDRFYADIADTSLGNDKRLATEAISAKKEFEKTK 413
Db 350 SKFLGSFRSFKLTNPANATVYKISDDEFVGSNSTVADADKYKAGDFLNEATADFCKQIQ 409
Qy 414 TIPAEVKTLEIPEK---PDKQPKQOESL 441
Db 410 ANPT-NYKSVLGIPTMLINDAKDNEKASL 439

```

## RESULT 2

0980L5 PRELIMINARY: PRT: 461 AA.

```

AC 0980L5;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE ABC TRANSPORTER XYLOSE-BINDING LIPOPROTEIN.
GN MYPU_3460.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CT1P;
RX MEDLINE=21267165; PubMed=11353084;
  Chambaud I., Hellig R., Ferris S., Barbe V., Samson D., Galisson F.,
  Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
  Blanchard A.;
  "The complete genome sequence of the murine respiratory pathogen
  Mycoplasma pulmonis.";
  Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL: AL445564; CAC13519.1; -.
DR Mypulast: MYPU_3460; -.
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp; 1.
KM Complete proteome.
SQ SEQUENCE 461 AA; 51096 MM; 6724D4D820809CE4 CRC64;

```

Query Match 24.1%; Score 554.5; DB 16; Length 461;  
 Best Local Similarity 30.3%; Pred. No. 4.1e-24;  
 Matches 140; Conservative 88; Mismatches 175; Indels 59; Gaps 13;

```

Qy 5 IKMNFLLGLGVLPPLSATATISAGCWDKETTKEKSDADNONKOIT--DVSKISGLVNER- 61
Db 1 KMLNKK-KLFSLLPYAALAIAPATF-----VSCAQNPKNKTNNLSDSKITTDLSQKE 51
Qy 62 -----KSEITMAKADANKHGLNMAIVTAGTVNDNSFNQSSWEAIIQOLGALTGGEI 113
Db 52 VTETQIKVENKIKQASLETFQK-----VVLITPADGNIDDKSFNQGYEESQKTLAKDPVDKAY 106
Qy 114 TS-----VDSSTALEGKYSLSLANTNKNRWVLSGFGHDAFTWLIKIPENK 159

```

```

Db 107 KSONREAEENOHLKDYNSAVKLEQNKVVALDRGYTTWLTGFOQGEIENFLDENDNL 166
Qy 160 QLFTEKNIIIGIDWTDTEEN-VIPTGRYINLYKTEEAGMLAGYANASFLAKKPPSDPTK 218
Db 167 RRFKEKRVKIIIVDAPNANNSKIIPQSLISLFLKTEEBAGMAGYASADFLGKRYANNEAK 226
Qy 219 RSAIVIGGISPATVDFIAGYLAGIRAMN--LKNSDKTKITTDKIEIINLGFVDVDTSTK 276
Db 227 RAISAFGGDPAVYVDLNLGFEFGIRAMNSAEANKNKRYIVSENLVLDTF-IPNAEKN 285
Qy 277 ERLEQIASDKRSTLLAVAGPTLEIFSDI--ANONDRYLIGVTDQSLVYTKTKNKF 333
Db 286 EVASNVETGKSTISLPAVAGPTGVVVDLARDTSDERFIVGVDTSLSFTNDSKRF 345
Qy 334 TSLIKNLGYVSFVSLDYT-----KSNRNLAGFEFGKSAATVYLGIKDRFVQIA 385
Db 346 TSVINAIAPVQIILLALTNDKDESVIILKEGDKFLGS---NPKMLVLKRLISAKFVNIT 402
Qy 386 DTSLEGNDRKRLATEAISEA--KKEFEKTKTIPAEVKTLE 425
Db 403 KSRVKEISIKTQADTSIQKAIKDMANPNNSKRIEKEMTNGDLE 444

```

## RESULT 3

09X775 PRELIMINARY: PRT: 465 AA.

```

AC 09X775;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE P48 MEMBRANE LIPOPROTEIN PRECURSOR.
GN P48.
OS Mycoplasma agalactiae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2110;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M7;
RX MEDLINE=20002620; PubMed=10531294;
  Rosati S., Pozzi S., Robino P., Montinaro B., Conti A., Fadda M.,
  Pittau M.;
  "P48 major surface antigen of Mycoplasma agalactiae is homolog to a
  malt product of Mycoplasma fermentans and belongs to a selected family
  of bacterial lipoproteins.";
  Infect. Immun. 67:6213-6216(1999).
DR EMBL: AJ132423; CAB43718.1; -.
DR InterPro: IPR003760; Bmp.
DR Pfam: PF02608; Bmp; 1.
KM Signal; Lipoprotein.
FT SIGNAL 1
FT CHAIN 22
FT 23 465 P48 MEMBRANE LIPOPROTEIN.
SQ SEQUENCE 465 AA; 51149 MM; 60AD5448CFE03C6 CRC64;

```

Query Match 18.1%; Score 417; DB 2; Length 465;  
 Best Local Similarity 28.4%; Pred. No. 3.1e-16;  
 Matches 141; Conservative 94; Mismatches 180; Indels 82; Gaps 22;

```

Qy 3 KIKKNNKFLGLVFPPLSATATISAGCWDKETTKEKSDADNONKOITDVSKISGL-VNER 61
Db 2 KKNKRYLFLGAA--PVLSPVLVAASCGDKYFKETEVGVGTIISTLHITSRKGLKLEEG 58
Qy 62 KSEITMAKADANKHGLNMAIVTAGTVNDNSFNQSSWEAIIQOLGALTGGEITSVDSSTA 121
Db 59 LTVENAPKA-----TTITDEGSVHDESFNQGAEVAKVSYELGDLKRAQV-SGNK 107
Qy 122 ELEGR-----YSLANTNKNRWVLSGFOHDAFTRWLIKIPENK-QLFETKNNII 168
Db 108 NLRNKVPYEPKKGQILLEAKNMIIDSGFRIVYLCGFHQASL---VGLDNIYIKKIKIDNNII 164
Qy 169 ILGIDWT-----DTENV-----IPTGRYINLYKTEEAGMLAGYANASFLAKKPPSDPTK 218

```



```

Db      165 FTTVDENFETEDDANKPIKIKIGEGHLYPVLTFRQKAAYINAGRALADYFSQVYKDOPK 224
Oy      219 KSAIVGGISPAVTDFEITAGYLGIKAMWLNKSKDKRTTDTCKIEINLGEVDYDSTKER 278
       | : ||| ||||| ||| ||| : ||| : ||| : ||| :
Db      225 RTGAFAGGIPMWAEITFIAGTFPGJILDMWKHEPKEAKTKSLNETIELNTLF---TSQPQA 281
Oy      279 LEQIASKDKPSLLAVAGFL-TEITSDI--IANONDRYLISGVDTQSGLVYTKN----- 330
       | | | | : | | | | : | : | : | : | | | | | | | |
Db      282 TTAISVVKAPASYSFAAGSLSTDTPAKEIKKLADK-DKFLTGVDADO-----KNALKGH 333
Oy      331 KEFTSILKNLGYSVHSVLSDLYTKKSNSHNL-AGEFEGRKSAT-VYLGIKD---REYDI 384
       | : | | : | : | : | : | : | : | : | : | : | : |
Db      334 RIFTSMVKRIQGAAYNIILADLSXSGENQDLDPGFEIGKKNGTPTFYVGGDIEDKQYGV 393
Oy      385 APTS-LEGDNKRLATEAISAKKEPEDEKKTTPABEVKTLDEL-----PEMPDKQ 433
       | : | : | : | : | : | : | : | : | : | : | : |
Db      394 ATSGILDLDNDINDIAKALMDATAYVOK-KT---EIQSLKDOMEYAKKALGARFPSPD 448
Oy      434 PDKQESLDDKLTITDINK 450
       | : | : | : |
Db      449 GGQFGKMVDWLMSLETRK 465
SUBMIT 4
ID      052311 PRELIMINARY; PRY; 428 AA.
AC      052311;
DT      01-JUN-1998 (TREMBLrel. 06, Created)
DT      01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      MEMBRANE LIPOPROTEIN P48V.
OS      Mycoplasma fermentans.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
CC      Mycoplasmataceae; Mycoplasma.
OX      NCBI_TextID=2115;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-PG18;
RA      Rawadi G., Dyer K., Dujeancourt A.;
RL      Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF036106; AAB89740.1; -
DR      InterPro: IPR003760; Bmp.
DR      Pfam: PF02608; Bmp; 1.
KW      Lipoprotein.
SO      SEQUENCE 428 AA; 47862 MW; D9506E817E330EEA CRC64;
```

```

Query Match      17.4%: Score 400.5; DB 2; Length 428;
Best Local Similarity 27.3%: Pred. No. 2.4e-15;
Matches 129; Conservative 95; Mismatches 175; Indels 73; Gaps 21.

QY      5 IKMNFPLGLGLVPLPSAI-ATISAGCMQKETTKEKRSADNOKOI--IDVKSICLVNER 61
      : : : : : | | | | : : : : : | : : : : : | : : : : :
Db      1 MKRSKKILGLT-SPIAAILPAVAVSC-----KNDNESNISPEKIDISKRYTTTNGANG 50
      : : : : : | | | | : : : : : | : : : : : | : : : : :

QY      62 KSEIMAAKADAKHHGGLNMAIYTAGTVYDNDSEFNOSWFAIQOLALJGEGELTSDVSSIA 121
      | : : : : | : : : : | : : : : | : : : : | : : : : |
Db      51 KOVVAKNAEL----LKLKPVLLITDGGKIDDKFSOSAFAPALKAINKQGIIEINSVEPS-S 104
      | : : : : | : : : : | : : : : | : : : : | : : : : |

QY      122 ELEGKYSILANTNNKVMVYLSGFQHDADFTBLKIPENKQLFTEKNNIILGIDMTDTEANI 181
      | : : : : | : : : : | : : : : | : : : : | : : : : |
Db      105 NFESAYNSALSGHKIHWLNGFKHOOOSIKOY--IDAHRELELRNOIKITIGIDE-DELEY 161
      | : : : : | : : : : | : : : : | : : : : | : : : : |

QY      182 PTGRYINILYTKTEEGWLAGVYANASFLAKFPSPDTKRAIYVIGGISPAYMDFTAGYLA 241
      : : : : : | : : : : | : : : : | : : : : | : : : : |
Db      162 KM--FYSLDFNFKESAFYTTGIALASWLSQ---DESKRVVASFGGAGPVGVTTFDEGFKK 216
      : : : : : | : : : : | : : : : | : : : : | : : : : |

QY      242 GIKANNLKNDSKDKTIT-TDKLIEINLGFVPDSTKERLEQIAS-----KDKPST 290
      | : : : : | : : : : | : : : : | : : : : | : : : : |
Db      217 GILYLYNQKH--KSVAFYTHSPYKLDGSG-----TAGEKNNYIYNWV,STPADYKYNPBY 269
      : : : : : | : : : : | : : : : | : : : : | : : : : |

QY      291 LIAVAGPLTEIFSDIILANONDRYLIGVDTDOSLVYTKTKNKEFTSII,KNLGYSVFSVSLD 350
      : : : : : | : : : : | : : : : | : : : : | : : : : |

```

```

Db      270 ILISVAGPAT--FEYKLANKGQYIYGVSDDGM--QDKRILTSVLEKHKQAVETFLD 325
Oy      351 I-----YTKSNSRNLTAGEFFGSKSATVYLGIIDRFVADIADTSLSENGDKL-ATE 399
Db      326 LILEEEGCKPYVVDKKADKKWSHFQ-----KEKIGVAENHFSFTEQAKINN 377
Oy      400 AISEAKKEFEETKTIIPAEVYRKLTLEIPMD-KQPKQKQESLDLKITDINK 450
Db      378 KIKEAKIKMFKK---LPEDFVKYIINSDKALDKGNKIDVSRLEFALISAIK 425

RESULT      5
G9RGX5
AC      G9RGX5      PRELIMINARY;      PRT:      428 AA.
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      MACOPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.
GN      MALP.
OS      Mycoplasma fermentans.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OX      Mycoplasmataceae; Mycoplasma.
RN      NCBI_TaxID=2115;
RP      [1]
RC      SEQUENCE FROM N.A.
RX      STRAIN-SK5;
RA      MEDLINE=99115554; Pubmed=9916088;
RA      Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RT      "Differential posttranslational processing confers intraspecies
RT      variation of a major surface lipoprotein and a macrophage-activating
RT      lipopeptide of Mycoplasma fermentans.";
RL      Infect. Immun. 67:760-771(1999).
DR      EMBL; AF099211; AADI635.1; -.
DR      InterPro; IPR003760; Bmp.
DR      Pfam; PF02608; Bmp; 1.
KW      Signal; Lipoprotein.
FT      SIGNAL      1      24      POTENTIAL.
FT      CHAIN      25      428      MACOPHAGE ACTIVATING LIPOPROTEIN-404.
SQ      SEQUENCE      428 AA;      47835 MW;      DD3F0F47EA2B1460 CRC64;

```

[illegible]

|        |   |  |                         |
|--------|---|--|-------------------------|
| Oy     | 402   | SEAKKEFEETKTPPAEVRKTEIPEMPD-KOPDKOESIDKLITDIK    | 450                     |
|        | :   | :   :  | :                       |
| Dd     | 380   | KKAINKFKE----LPEDFVATINDSKALDKDNKDIVSERLAITSAINK | 425                     |
| RESULT | 6   |  |                         |
| O32417 |   |  |                         |
| ID     | O32417  | PRELIMINARY;                                     | PRT; 428 AA.            |
| AC     | O32417;   |  |                         |
| DT     | 01-JAN-1998                                       | (TREMBREL. 05,                                   | Created)                |
| DT     | 01-JAN-1998                                       | (TREMBREL. 05,                                   | Last sequence update)   |
| DT     | 01-JUN-2001                                       | (TREMBREL. 17,                                   | Last annotation update) |
| DE     | M161AG.   |  |                         |
| OS     | Mycoplasma fermentans.                            |  |                         |
| OC     | Bacteria; Firmicutes; Bacillus/Clostridium group; |  | Mollitcetes             |
| CC     | Mycoplasmataceae; Mycoplasma.                     |  |                         |
| OK     | NCBI-TaxID=2115;                                  |  |                         |

SEQUENCE FROM N.A.  
TISSUE=BONE MARROW;  
MEDLINE=98022661; PubMed=9359703;  
RX Matsunoto M., Takeda J., Inoue N., Hara T., Hatnataka M., Takahashi K.,  
RA Nagasawa S., Aredo H., Seya T.;  
RA "A novel protein that participates in nonself discrimination of  
RT malignant cells by homologous complement.";  
RT Nat. Med. 3:1266-1270(1997).  
RL [2]  
RN  
RP  
RP  
RP Matsunoto M., Seya T.;  
RL Submitted (Apr-1999) to the EMBL/Genbank/DBD databases.

RP SEQUENCE FROM N.A.  
RX MEDLINE=98241611; Pubmed=9575156;  
RA Matsumoto M., Nishiguchi M., Kikawa S., Nishimura H., Nagasawa S.,  
RA Seya T.;  
RT "Structural and functional properties of complement-activating protein  
RT M16Ag, a Mycoplasma fermentans gene product that induces cytokine  
RT production by human monocytes.";  
RL J. Biol. Chem. 273:12407-12414(1998).  
RN [4]  
RN SEQUENCE FROM N.A.  
RP Seya T., Matsumoto M.;  
RA "MEMBRANOUS PROTEIN M16Ag AND CYCLIC-DNA CAPABLE OF CODING THE  
RT SAME.";  
RT  
RL Patent number JP1997157295 A/1,  
RL 17-JUN-1997, KAGAKI GIYUO SHINKO JIGYODAN.  
EMBL, D64083: BAA2530.1; -.  
DR EMBL, AB026157; BAA77211.2; -.  
DR InterPro, IPR005760; Bmp.  
DR Pfam, PF02608; Bmp; 1.  
SQ SEQUENCE 428 AA; 47862 MW; F43B078F21DAD6E CRC64;

|                       |                  |                   |           |            |
|-----------------------|------------------|-------------------|-----------|------------|
| Query Match           | 17.2%            | Score 394.5       | DB 2      | Length 428 |
| Best Local Similarity | 26.8%            | Pred. No. 5.4e-15 |           |            |
| Matches 126           | Conservative 101 | Mismatches 174    | Indels 69 | Gaps 20    |

[illegible][illegible]

|                  |   |              |              |
|------------------|---|--------------|--------------|
| RESULT           | 7   |              |              |
| Q9RGX6           |   |              |              |
| ID               | Q9RGX6  | PRELIMINARY; | PRT; 429 AA. |
| AC               | Q9RGX6;   |              |              |
| DT               | 01-MAY-2000 (TREMBLrel, 13, Created)                          |              |              |
| DT               | 01-MAY-2000 (TREMBLrel, 13, Last sequence update)             |              |              |
| DT               | 01-DEC-2001 (TREMBLrel, 19, Last annotation update)           |              |              |
| DE               | MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.              |              |              |
| GN               | MALP.   |              |              |
| OS               | Mycoplasma incognitus.  |              |              |
| OC               | Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes; |              |              |
| CC               | Mycoplasmataceae; Mycoplasma.                                 |              |              |
| NCBI_TaxID=2101; |   |              |              |
| OX               |   |              |              |

|    |  |         |           |                                       |
|----|--|---------|-----------|---------------------------------------|
| RP | SEQUENCE FROM N.A.   |         |           |                                       |
| RC | STRAIN-INCOGNITUS;   |         |           |                                       |
| RX | MEDLINE-99115554; PubMed-9916088;                                    |         |           |                                       |
| RA | Calcutt M.J., Klim M.F., Kaipras A.B., Muhrirad P.F., Wise K.S.;     |         |           |                                       |
| RT | "Differential posttranslational processing confers intraspecies      |         |           |                                       |
| RT | variation of a major surface lipoprotein and a macrophage-activating |         |           |                                       |
| RT | lipopeptide of Mycoplasma fermentans.";                              |         |           |                                       |
| RL | Infect. Immun. 67:760-771(1999).                                     |         |           |                                       |
| DR | EMBL; AF099210; AAD16594.1; --                                       |         |           |                                       |
| DR | InterPro: IPR003760; Bmp.  |         |           |                                       |
| DR | Pfam: PF02608; Bmp; 1.   |         |           |                                       |
| KW | Signal; Lipoprotein.   |         |           |                                       |
| FT | SIGNAL   |         |           |                                       |
| FT | CHAIN  | 1       | 24        | POTENTIAL.                            |
| SO | SEQUENCE   | 429 AA; | 47961 MW; | MACROPHAGE ACTIVATING LIPOPROTEIN-404 |
|    |  |         |           | B6D09A8B12AC3171 CR664;               |

|                           |       |                    |            |             |
|---------------------------|-------|--------------------|------------|-------------|
| Query Match               | 17.1% | Score 394;         | DB 2;      | Length 429; |
| Best Local Similarity     | 27.0% | Pred. No. 5.8e-15; |            |             |
| Matches 128; Conservative | 101;  | Mismatches 169;    | Indels 76; | Gaps 21;    |

```

OY 5 IKMKNFGLGVLPFLSL-ATISAGCDMKETTRKESADONOKRITVSKISGLVNERKS 63
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKRSKRITLGL-SPIMALPRVAVSCGNDES-----NISFKERISKITTANNANGKO 52

OY 64 EIMAKADANKHFLGNMAIYTAGTIVDNDSFNOSMWEAIOQLGALVGETTIVSDSSTAEL 123
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 53 VYKNAEL-----LTKLEPVLLITDESKIDKSDKSNOSAFALKAINKQTSIEINNVEPS-SNF 106

OY 124 EGRYSSLANTKKNVWVLSFGDHGDAFTRMLKIPBNKOLFTEKNITLIGIDMTDENTPIPT 183
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 ESAYNSALSAGCHKIWLVLNGFRHOOSIOY--IDAHRELEBERNOKITLIGIDP-DIEIEYKW 163

OY 184 GRYINTLYTKEEAGWLAGVANASFIARKFPSPDPRKSAIYIGGISPBAVDTFIAGYLAGI 243
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 164 --FYSLOFINIKESAFITGYALIASWLSBQ--DESKRVVASFGGAFGPGVTFNNGEFAKGI 218

OY 244 KAMLUKMSDKKTKI-TYDKIEINIGFVODTSTYERLEQJAS-----KDKPSTLL 293
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 219 LYYOKH--SSKRIYHSPVAKDSGF-----TAEKKNVYINNVLSSTPADVKNPVIIL 277

```

[illegible]

```

Query Match      17.1%: Score 394; DB 2; Length 429;
Best Local Similarity 27.0%: Pred. No. 5.8e-15;
Matches 127; Conservative 102; Mismatches 172; Indels 70; Gaps 21.

0Y      5 IKMKNFELGLVFLPSAI-ATISACCMCKETTKREKSDADNOKQITIV)SKISGLVNERKS 63
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MKRSKKILLGL-SPITAILPAAVASCNNDES-----NISFKERDI)SKYTTTANAKQ 52

0Y      64 EIMAKKADANHFGLMNAIVTAGGVNDNSFNQSSWEAICQIGALITGEGITVSDSSTAE 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
53 VYKNEL-----LKLKPVLLIDEGCIDDKSFNQSAFELKAI)NKQIT)EINNVEPS-SNF 106

0Y      124 EGGYSSILANTKNNVWLISGFQHDGAF)RWLKLPEKCOLFTEKNIITILGIDMTDEV)IPT 183
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
107 ESATVSALSAGHAKIVLWLGFKHQOSIKQY--IDAREIELERNOIKITGIDF-DILEYKW 163

0Y      184 GRVILVYTKTEBAGWLAGVANA)SPFLAKKFPSPDTRSAIVIGSGISPA)TDEITAGI)AGI 243
      : : : : : : : : : : : : | : | : | : | : | : | : | : | : | : | : |
184 --FYSLOENIKESAFTTGYAIA)SWLSEQ---DESKRVVASFEGGAFPGVTTTNEGEFAKTI 218

0Y      244 KAMNLKNSDDKKRTI-TTQKIEI)NLF)FDVOD)STKRELEOIAS-----KDRPSTLL 292
      : : | : | : | : : : : : : : : : : : : : : : : : : : : : : : : : :
219 LYINQKH--KSKIIYHTSPVKIKD)SGF-----TAGKEMQV)INNVLS)TPADVAKY)NPHVLL 271

0Y      293 AVAGPLEIFISDI-IANONDRYLIGVD)DQSLVYTKTKNKFET)IL(KLGYSVFV)SLDL 351
      : : | : | : | : : : : : | : | : | : | : | : | : | : | : | : | : |
272 SVAGPAT--FETVRLAANKGGYIVG)VDSPQGM-I--QDKRRLITV)YLKHK)KAVYETI)LDL 327

0Y      352 -----YTKKNSNRNL)AGFEFGK)SAIVYLCIKD)RFVDIADT)S)LEGNDKTL-ATEA 400

```

| Db     | 338                             | ILEKEGKPPVAVD                           | DKDKAKWMSHFQI           | -----      | KEKMGVAENHFS | TEEDAKI          | LNK        | 379              |     |     |
|--------|---------------------------------|---|-------------------------|------------|--------------|------------------|------------|------------------|-----|-----|
| Qy     | 401                             | ISEAKKEEKT                              | KTTPAEV                 | KRTLEI     | EMPD         | -KOPDR           | QOESL      | DLITDK           | 450 |     |
|        |                                 |   | :                       | :          | :            | :                | :          | :                | 111 |     |
|        |                                 |   | :                       | :          | :            | :                | :          | :                | 111 |     |
| Db     | 380                             | IKKAIKKME                               | KE---                   | LPEDFV     | ATINSD       | AKKGN            | KNDIV      | SESLKAI          | LNK | 426 |
| RESULT | 9                               |   |                         |            |              |                  |            |                  |     |     |
| O9RGX7 |                                 |   |                         |            |              |                  |            |                  |     |     |
| ID     | O9RGX7                          |   | PRELIMINARY;            |            |              | PRT;             | 428        | AA.              |     |     |
| AD     | O9RGX7                          |   |                         |            |              |                  |            |                  |     |     |
| DT     | 01-MAY-2000                     | (TREMBLrel. 13,                         | Created)                |            |              |                  |            |                  |     |     |
| DT     | 01-May-2000                     | (TREMBLrel. 13,                         | Last sequence update)   |            |              |                  |            |                  |     |     |
| DT     | 01-JUN-2001                     | (TREMBLrel. 17,                         | Last annotation update) |            |              |                  |            |                  |     |     |
| DE     | MACROPHAGE                      | ACTIVATING                              | LIPOPROTEIN-404         | PRECURSOR. |              |                  |            |                  |     |     |
| GN     | MALP.                           |   |                         |            |              |                  |            |                  |     |     |
| OS     | Mycoplasma                      | fermentans.                             |                         |            |              |                  |            |                  |     |     |
| OC     | Bacteria;                       | Firmicutes;                             | Bacillus/Clostridium    | group;     | Mollicutes;  |                  |            |                  |     |     |
| OC     | Mycoplasmataceae;               | Mycoclasma.                             |                         |            |              |                  |            |                  |     |     |
| OX     | NCBI_TaxID=2115;                |   |                         |            |              |                  |            |                  |     |     |
| RN     | [1]                             |   |                         |            |              |                  |            |                  |     |     |
| RP     | SEQUENCE FROM N.A.              |   |                         |            |              |                  |            |                  |     |     |
| RC     | STRAIN-II-29/1;                 |   |                         |            |              |                  |            |                  |     |     |
| RC     | MEDLINE=9911554;                | pubmed=9916088;                         |                         |            |              |                  |            |                  |     |     |
| RA     | Calcutt M.J., Kim M.F.,         | Karpas A.B., Muhlradt P.F.,             | Wise R.S.;              |            |              |                  |            |                  |     |     |
| RT     | "Differential posttranslational | processing confers intraspecies         |                         |            |              |                  |            |                  |     |     |
| RT     | variation of a major surface    | lipoprotein and a macrophage-activating |                         |            |              |                  |            |                  |     |     |
| RL     | Lipopeptide of Mycoplasma       | fermentans.";                           |                         |            |              |                  |            |                  |     |     |
| RL     | Infect. Immun. 67:760-771       | (1999).                                 |                         |            |              |                  |            |                  |     |     |
| DR     | EMBL; AF099209; AADI6333.1;     | --.                                     |                         |            |              |                  |            |                  |     |     |
| DR     | InterPro; IPR003760; Bmp.       |   |                         |            |              |                  |            |                  |     |     |
| DR     | Pfam; PF02608; Bmp. 1.          |   |                         |            |              |                  |            |                  |     |     |
| KW     | Signal: Lipoprotein.            |   |                         |            |              |                  |            |                  |     |     |
| FT     | SIGNAL                          | 1                                       |                         |            |              |                  |            |                  |     |     |
| FT     | CHAIN                           | 25                                      | 428                     |            |              | POTENTIAL.       |            |                  |     |     |
| SO     | SEQUENCE                        | 428                                     | AA;                     | 47960      | MM;          | D6B69C0BA969CEP1 | CRC64;     |                  |     |     |
|        |                                 |   |                         |            |              | MACROPHAGE       | ACTIVATING | LIPOPROTEIN-404. |     |     |

```

17.1%; Score 393.5; DB 2: Length 428;
Query Match Similarity 26.8%; Pred. No. 6.2e-15;
Best Local Similarity 26.8%; Pred. No. 6.2e-15;
Matches 127; Conservative 97; Mismatches 178; Indels 71; Gaps 21.

QY 1 MKKKIKMKNFKLGLGVPLPISATATISAGCWDKETEKEKSADNONKQITDVSKISGLVNE 60
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1 MKKSKK--TLGLSPILAILPAPVANS CG-----NNDESNISFKEDIDIKYTTTN--ANE 50

QY 61 KKSEITMAAKDANKHPEGLMMAIVTAAAGVYNDNSFNQSSNEAIIQULGALTGGEITVS DST 120
   ::::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::
Db 51 KQ--VKNNAELIK--LKPVLITDEGKIDDKSPNQAFAEALKAIKQTCIEINNVEPS- 103

QY 121 ALEBEKYSLSLANTNNKNMVVLTSGFHGDAFTRKLKIPENKQUTTEKNITIIILGDMWDTLEV 180
   ::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 104 SNEEASVANSALSAGHKIWLNLGFKHQOQSIKQY--IDAHEELERNOIKITIGDF--DIEFE 160

QY 181 IPTGRIYNTLTKTEEAGWLAGYANASFLAKKRPSPDTRKSAIVAGGISPAVTFDEIAGYL 240
   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db 161 YKM--FYSGIFPIKRSAPFTGTGAILMSWSEQ---DESKRVVAPSPGGAGFPGYTTTNEGGA 215

QY 241 AGIKAMNLKNSDKTKRI--TTDKIEINLGLVDYDTSTKERLEQIAS-----KDKPS 289
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 216 KGIILYNNOKH--KSKSIYHTSPVKIDSGF-----TAGEEKMNTVINNVLSSTPADVKYNPH 268

QY 290 TLLVAVGLPTEIFSDIINANDRYLILGYDQDSLYYTKKKNFFSILKNLGIYSFVSYS 349
   ::|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 269 VILSVAGPAT--FETVRILANKGQVYIAGVSDGMI--QDKRILITSVLKHIQAQVETLL 324

QY 350 DL-----YTKKSNRSRLAGECFGKRSATVYLGIDRVDIADITSLEGNDKYL--AT 398
   ::::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::  :::
Db 325 DLILEKEGKYRYVVKDKKADKKKNSHFGIQ-----KKNIGAGVLENNFSSTEGDAKIN 376

QY 399 EAIISAKKEFEETKTYTIPAEVYRKTLEIPEMD--KQPKDQOESDLKLTITDINK 450
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

DB 377 NKIKAEIMFKR-----LPEDFVKYINSDKALKDGNKIDNVSERLEAIIISAINK 425

RESULT 10  
09RGX3 PRELIMINARY; PRT; 428 AA.

AC 09RGX3:  
DB 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE MACROPHAGE ACTIVATING LIPOPROTEIN-404 PRECURSOR.  
GN MALP.  
OS Mycoplasma fermentans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2115;  
RN [1]  
RP SEQUENCE FROM N.A.  
RT STRAIN-K7.  
MEDLINE=99115554; PubMed=9916088;  
RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;  
RT "Differential posttranslational processing confers intraspecies  
variation of a major surface lipoprotein and a macrophage-activating  
lipopeptide of Mycoplasma fermentans.";  
RL Infect. Immun. 67:760-771(1999).  
DR EMBL; AF099213; AAD16397.1; -.  
DR InterPro; IPR003760; Bmp.  
DR Pfam; PF02608; Bmp; 1.  
KW Signal; Lipoprotein.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 428 MACROPHAGE ACTIVATING LIPOPROTEIN-404.  
SQ SEQUENCE 428 AA; 47875 MW; C7A536B409A60132 CnC64;

Query Match 16.9%; Score 389.5; DB 2; Length 428;  
Best Local Similarity 26.8%; Pred. No. 1e-14;  
Matches 126; Conservative 99; Mismatches 176; Indels 69; Gaps 20;

DB 5 IKWNFELGLVFLPSAI-ATISACMDKETTKEKSADNONKQITDVSKISGLVNERKS 63  
1 MKSKKILGL-SPIAALPAVAVSCGNDES-----NISFEKDISKYTTNANGK 52  
DB 64 EIMAKADANHFGLMAIVTAGTVNDNSFNQSSWEAIOQLGALTGEITSVDSSTAE 123  
53 VKNAEL-----LKLKPLVITDEKIDKSFNQSFAELKAIKNOTGIEINSVEPS-SNF 106  
DB 124 EGYSSSLANTKKNVYLSGFGHDAFTRLMKIPENKOLFTEKNIILIGIDMTDTENVIPT 183  
107 ESAVNSALSAGHKIMVNLGFKHQOSIKOY--IDAHELEENNOIITIGIDF-DIETEKW 163  
DB 184 GRYINLYKTEBAGWLAGYANASFLAKKFPSPDPTKRSIVIGGIGSPAVTDFIAGYLAGI 243  
164 --FYSLOFNKESAFETTGVAIASWLSQ---DESKRVVASFGVGFPGVTTFNEGFAGGI 218  
DB 244 KAMNKNSDKRTKI-TTQKIEINLGFVDVDTSTKERLEQIAS-----KDKPSTLL 292  
219 LYYNOKH--KSKSIYHTSPVKLDSG-----TAGKMTVYINNVLSSPADVKYPNHYIL 271  
DB 293 AVAGPLEIFSDIIANQDRYLIGVDQSLVYTKNKFETSLKNGSVFVSLDL- 351  
272 SVAGPAT--FETVRLANKGQYVIGVSDQGM--QDKRILITSVLKHKQAVETFLDLI 327  
DB 352 -----YTKKSNRSLAFEGGKKSATYVYLGIDKRFVADIADTSLGDNKRL-ATEAI 401  
328 LEKEBGYPYVVKDKKAKKWSHFQO-----KEKWIGVAENHFSNTEGQAANKNKI 379  
DB 402 SEAKKEFEKRTKTPAEVVRTLEIPEMPD-KOPDKQOESLDKLTITDINK 450  
380 KEAIKMFKE---LPEDFVKYINSDKALKDGNKIDNVSERLEAIIISAINK 425

RESULT 11  
09R3N6

ID 09R3N6 PRELIMINARY; PRT; 428 AA.

AC 09R3N6:  
DB 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE LIPOPROTEIN MALP-404 PRECURSOR.  
GN MALP.  
OS Mycoplasma fermentans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2115;  
RN [1]  
RP SEQUENCE FROM N.A.  
RT STRAIN-PG18, M39A, AND M70B;  
RC MEDLINE=99115554; PubMed=9916088;  
RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;  
RT "Differential posttranslational processing confers intraspecies  
variation of a major surface lipoprotein and a macrophage-activating  
lipopeptide of Mycoplasma fermentans.";  
RL Infect. Immun. 67:760-771(1999).  
DR EMBL; AF100324; AAD25736.1; -.  
DR EMBL; AF099214; AAD16398.1; -.  
DR EMBL; AF099215; AAD16399.1; -.  
DR InterPro; IPR003760; Bmp.  
DR Pfam; PF02608; Bmp; 1.  
KW Signal; Lipoprotein.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 38 MACROPHAGE ACTIVATING LIPOPEPTIDE MALP-2.  
SQ SEQUENCE 428 AA; 47863 MW; 59808324E218F03C CnC64;

Query Match 16.9%; Score 387.5; DB 2; Length 428;  
Best Local Similarity 26.6%; Pred. No. 1.4e-14;  
Matches 125; Conservative 101; Mismatches 175; Indels 69; Gaps 20;

DB 5 IKWNFELGLVFLPSAI-ATISACMDKETTKEKSADNONKQITDVSKISGLVNERKS 63  
1 MKSKKILGL-SPIAALPAVAVSCGNDES-----NISFEKDISKYTTNANGK 52  
DB 64 EIMAKADANHFGLMAIVTAGTVNDNSFNQSSWEAIOQLGALTGEITSVDSSTAE 123  
53 VKNAEL-----LKLKPLVITDEKIDKSFNQSFAELKAIKNOTGIEINSVEPS-SNF 106  
DB 124 EGYSSSLANTKKNVYLSGFGHDAFTRLMKIPENKOLFTEKNIILIGIDMTDTENVIPT 183  
107 ESAVNSALSAGHKIMVNLGFKHQOSIKOY--IDAHELEENNOIITIGIDF-DIETEKW 163  
DB 184 GRYINLYKTEBAGWLAGYANASFLAKKFPSPDPTKRSIVIGGIGSPAVTDFIAGYLAGI 243  
164 --FYSLOFNKESAFETTGVAIASWLSQ---DESKRVVASFGVGFPGVTTFNEGFAGGI 218  
DB 244 KAMNKNSDKRTKI-TTQKIEINLGFVDVDTSTKERLEQIAS-----KDKPSTLL 292  
219 LYYNOKH--KSKSIYHTSPVKLDSG-----TAGKMTVYINNVLSSPADVKYPNHYIL 271  
DB 293 AVAGPLEIFSDIIANQDRYLIGVDQSLVYTKNKFETSLKNGSVFVSLDL- 351  
272 SVAGPAT--FETVRLANKGQYVIGVSDQGM--QDKRILITSVLKHKQAVETFLDLI 327  
DB 352 -----YTKKSNRSLAFEGGKKSATYVYLGIDKRFVADIADTSLGDNKRL-ATEAI 401  
328 LEKEBGYPYVVKDKKAKKWSHFQO-----KEKWIGVAENHFSNTEGQAANKNKI 379  
DB 402 SEAKKEFEKRTKTPAEVVRTLEIPEMPD-KOPDKQOESLDKLTITDINK 450  
380 KEAIKMFKE---LPEDFVKYINSDKALKDGNKIDNVSERLEAIIISAINK 425

RESULT 12  
0992H4  
ID 0992H4 PRELIMINARY; PRT; 350 AA.  
AC 0992H4:  
DB 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, last sequence update)  
 DE PUTATIVE LIPOPROTEIN. 19, last annotation update)  
 DE PUTATIVE LIPOPROTEIN.  
 GN SPY128.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 NCBI\_TaxID=1314;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;  
 RA MEDLINE-21192684; Pubmed-11296296;  
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lal H.S., Lin S.P.,  
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
 RL EMBL; AEO06563; AAK34087.1; -;  
 DR EMBL; AEO06563; AAK34087.1; -;  
 DR InterPro: IPR001114; Adenylsuc-synt.  
 DR InterPro: IPR003760; Bmp.  
 DR InterPro: IPR001761; Peripla\_BP\_Like.  
 DR InterPro: IPR000217; Tubulin.  
 DR Pfam: PF02608; Bmp; 1.  
 DR Pfam: PF00532; Peripla\_BP\_Like; 1.  
 DR PROSITE: PS00227; TUBULIN; UNKNOWN\_1.  
 KW Lipoprotein; Complete proteome.  
 SQ SEQUENCE 350 AA; 36396 MW; 7B5A6A95E0BFBD2 CRC64;

Query Match 10.1%; Score 232.5; DB 16; Length 350;  
 Best Local Similarity 26.3%; Pred. No. 7.7e-06;  
 Matches 113; Conservative 57; Mismatches 153; Indels 107; Gaps 20;

QY 9 KFLGLVPLSLATIS-AGCWDKERTKEKSDNONKQITDYSKISGLVNEKSEIMA 67  
 DB 4 KFLGLG-----LASVAVLSLAACGRG-----A 26  
 QY 68 AKADANKHFLGMAIVTAGTVNDNSFNOSWEAIOQLGALTG-GETTSVDSSTAELEGK 126  
 DB 27 SKGASGSKTLKAMVYDITGVDKSNQASMEGLQSGKEMGKRTGTGDFPSTSESE 86  
 QY 127 YSLSLATNRKVVWVLSGFQ--HGDAFTRWLKIPEKNTLIIIGIDMTDENVIPTG 184  
 DB 87 YA-----TSLDTAVSGGQILYIGIFA--LKDALAKAGDNEGVFYIID--DIIRGD 136  
 QY 185 RYINLYKTEEAGMLAGYANASFLAKFPSPDKRSAL-VIGGISPAVDFITGIYAGI 243  
 DB 137 NVASVTFPADHEAAYLAGIAAK-----TTKTKTVGFVVGMEGTVITREKGFPAV 187  
 DB 244 KANMLKNSDKRTKITTDKIEINLGF---DVODSTKERLEQLASKDKPSTLLAVAGPL 298  
 DB 188 KS-----VDDTIQVYDYAGSPGDAKGTIAAQAAGAD--VIYAAAGTG 233  
 QY 299 TELPSDIIANONDR-----YLIGVDTQ--SLVYT---KTKNKFSTILKNGSVFS 346  
 DB 234 AGVFNEKAKEINERSEADKTVWVIGVDRDQDEKGYTSKDEKAEAFYLAASIKVEKAVOL 293  
 QY 347 VLSDLTTRKNSRNLAGFERGKKSATVYLGIDRFVDIADTSLEGDKRLATEISAKK 406  
 DB 294 INKQVADKK-----FPGK--TVY-GLKDGVEIATF---NVSKEAVKAIKEKA 338  
 QY 407 EFEKTKTIP 416  
 DB 339 KIKSGDIKVP 348

RESULT 13  
 Q9EV91 PRELIMINARY; PRT; 349 AA.  
 AC Q9EV91;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE PUTATIVE SUBSTRATE BINDING LIPOPROTEIN PRECURSOR OF AN ABC  
 DE TRANSPORTER.  
 GN BMDA.  
 OS Clostridium sticklandii.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 OC Clostridium.  
 NCBI\_TaxID=1511;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM 519T.  
 RA MEDLINE-21089007; Pubmed-11271425;  
 RA Graentzdorffer A., Plich A., Andreessen J.R.;  
 RT "Molecular analysis of the grd-operon encoded proteins of the glycine  
 RT reductase and thioresdoxinsystem from Clostridium sticklandii.";  
 RT Arch. Microbiol. 175:8-18(2001).  
 RL EMBL; AJ276209; CAC14304.1; -;  
 DR EMBL; AJ276209; CAC14304.1; -;  
 DR InterPro: IPR003760; Bmp.  
 DR Pfam: PF02608; Bmp; 1.  
 KW Lipoprotein.  
 SQ SEQUENCE 349 AA; 37083 MW; 53AE45114FCD5B2E CRC64;

Query Match 10.0%; Score 229.5; DB 2; Length 349;  
 Best Local Similarity 25.6%; Pred. No. 1.1e-05;  
 Matches 112; Conservative 66; Mismatches 145; Indels 115; Gaps 22;

QY 1 MKKKIKNNKFLGLVPLSLATISAGCWDKERTKEKSDNONKQITDYSKISGLVNE 60  
 DB 1 MKKRI---TALFLASINGIMTLACSSA--PAETPAEPAE-----E 38  
 QY 61 KRSIMAKADANKHFLGMAIVTAGTVNDNSFNOSWEAIOQLGALTGGETTSVDS-S 119  
 DB 39 TAAE-----SLKISMVTVGVKQDSNQSAMELEKAKKDLGLEIETIESKQ 86  
 QY 120 TAELEGYSILAN-TKNVWVLSGFQHGDAFTRWLKIPEKNTLIIIGIDMTDF 177  
 DB 87 DADYEPNLETLVDENDLIMGV-GFKMDAI---LAAAGN---YPEQYATIDNDYADAT 139  
 QY 178 -ENVIPIGRYINLYKTEEAGMLAGYANASFLAKFPSPDKRSAL-VIGGISPAVDF 235  
 DB 140 PDNV-----LGVLFKEEPSYLVG---LIAGM---TQSNKIGTIGMDVPVINF 184  
 QY 236 IAGYIAGIKANMLKNSDKRTKITTDKIEINLGFVODSTKERLEQLASKDKPSTLLAVA 295  
 DB 185 RFGFLAGYAKAN-----PDVAV---VDVQ-----FYNAFNDPAKGRAYA 219  
 QY 296 GPLTEIFSDII-----ANONDRYLIGVDTQSLVYTKRNFSTILNGLG 341  
 DB 220 NQMTASGVDTLFFHAAGDSNGAIEPAKEQNKNWVIGVDRDQ---DLAPDNVITSAYKRV 276  
 QY 342 YSVFVSLDLYTKSNSRNLAGFERGKKSATVYLGIDRFVDIADTSLEGDKRLATEAI 401  
 DB 277 NAMFNVAKEI-----KQGNFAG-----GTTIVGLAEGVDIAPYTSK-NVPADILEFV 324  
 QY 402 SEAKKEFEKTKTIPAE 419  
 DB 325 EAEKQKIINKETIYVGT 342

RESULT 14  
 Q97RHO PRELIMINARY; PRT; 350 AA.  
 AC Q97RHO;  
 DT 01-OCT-2001 (Tremblrel. 18, Created)  
 DT 01-OCT-2001 (Tremblrel. 18, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE LIPOPROTEIN.  
 GN SP0845.  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 NCBI\_TaxID=1313;

[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-TIGR4;  
RC MEDLINE-21357209; PubMed-11463916;  
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,  
Durkin A.S., Gwin M., Kolony J.F., Nelson W.C., Peterson J.D.,  
Rumaym L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,  
McDonald L.A., Feldlyum T.V., Anglucci S., Dickson T., Hickey E.K.,  
Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
R.A. Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.,  
"Complete genome sequence of a virulent isolate of Streptococcus  
RT pneumoniae";  
RL Science 293:498-506(2001).  
DR EMBL: AE007390; AAK74976.1; -  
DR TIGR: SP0845; -  
DR InterPro: IPR001114; Adenylsucc\_synth.  
DR InterPro: IPR003760; Bmp.  
DR InterPro: IPR00217; Tubulin.  
DR Pfam: PF02608; Bmp; 1.  
DR PROSITE: PS00227; TUBULIN; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 350 AA; 36746 MW; CBD767BFF501E69 CRC64;

Query Match 9.7%; Score 223; DB 16; Length 350;  
Best Local Similarity 24.8%; Pred. No. 2.7e-05;

Matches 110; Conservative 63; Mismatches 146; Indels 122; Gaps 22;

OY 1 MKKTKMKNFGLGLVPLSAITATIS--ACGMDKETTKEKSADNOKQITDVSKISGLVN 59  
DB 1 MKKR-----CMGLGLVPLSAITATIS--ACGMDKETTKEKSADNOKQITDVSKISGLVN 59  
OY 60 EKSEIMAKKADANKHFGMLNMAIVTAGTVNDNSFNQSMGALQOLGA---LTGGETITSV 116  
DB 40 -----AAIVTDGVDKSFNQSMEGLQAMGKSHNLKONGFTY 79  
OY 117 DSTAELE-----GKYSILANTNKNVWVLSGFGHDGAFTRMLKIPENKOLFTEKNI 167  
DB 80 FQSTSEADVANNLQQAAGSYNLIFFGV-----GFLANNA-----VADAKKEHTDLNV 125  
OY 168 IILGIDMTDTEVNIPTGRITNLTYTEBAGWLAGYANASFLAKKPPSDPTKSAIYIGG 227  
DB 126 VALDIVIRDKQNV-----ASVTFADNESGYLAGVAAAK-----TTKRYQVFFVG 170  
OY 228 I-SPATDFEILAGIKAMNLSKSDKTKITTDKIEINLGFVDPTSKERLEQIASKD 286  
DB 171 IESVVISREAGFKAGV-----ASVDPSIKVQVDYAG-SFG-DAKKGKTIIAAQYAAAGD 223  
OY 287 KPSTLLAVAGPL-TEIFSDIIANONDR-----YLIGVDTQSL--VYT---KTKNKEF 333  
DB 224 ---IYYQVAGTGAGVFAEAKSLNRSRENEKRWYVIGVDRDQAEKGKYSKDGKESNFVL 280  
OY 334 TSLINKLGVSFVSLDLYTKKSNRNLGAFEPGKKSATVYLGINDRPDIADTSLSEGD 393  
DB 281 VSTLKQVGTIVDI-----SNKAERGEFPGGQ---VIYVSLDKGVDLAVTNLSEGG 329  
OY 394 KRLATEAISEAKKEFEKTKTIP 416  
DB 330 KR-----AVEDAKAKILDSGVKVP 348

RESULT 15

O9CFM9 PRELIMINARY; PRT: 350 AA.  
AC O9CFM9;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)  
DE BASIC MEMBRANE PROTEIN A.  
GN BMA.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Lactococcus.  
OX NCBI\_TaxId=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TL1403;  
RA MEDLINE-21235186; PubMed-11337471;  
RA Bolotin A., Wincker P., Manger S., Jallion O., Malarne K.,  
Weissenbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
RT lactis ssp. lactis TL1403";  
RL Genome Res. 11:731-753(2001).  
DR EMBL: AE006374; AAK0536.1; -  
DR InterPro: IPR003760; Bmp.  
DR Pfam: PF02608; Bmp; 1.  
KW Complete proteome.  
SQ SEQUENCE 350 AA; 36652 MW; 2A430D8EEDD4C814 CRC64;

Query Match 9.0%; Score 207; DB 16; Length 350;  
Best Local Similarity 24.2%; Pred. No. 0.00022;

Matches 107; Conservative 62; Mismatches 156; Indels 118; Gaps 20;

OY 1 MKKTKMKNFGLGLVPLSAITATIS--ACGMDKETTKEKSADNOKQITDVSKISG 56  
DB 1 MKKR-----VIAVSALALASVAVLACCRSHDAAGSGK----- 32  
OY 57 LVNERSEIMAKKADANKHFGMLNMAIVTAGTVNDNSFNQSMGALQOLGALTG-----G 111  
DB 33 -----AKTD-----LKAIVTEIGVNDNSFNQSMGALQSGKGNLKKGTG 75  
OY 112 EITSVDSSTAELEGKYSILANTNKNVWVLSGFGHDGAFTRMLKIPENKOLFTEKNIILIG 171  
DB 76 YTFQSNASASDPTTVNNSAEQGYKLLFGTGSIDATSAAK---NN---PKSNFVIYD 129  
OY 172 IMTDTENYIPTGRITNLTYTEBAGWLAGYANASFLAKKPPSDPTKSAI-VIGGCTSP 230  
DB 130 SVIKDQNV-----ASATFADNESAYLAGVAAK-----ATKTKKIGFIGQMOSD 174  
OY 231 AVTDFAGIAGIKAMNLSKSDKTKITTDKIEINLGFVDPTSKERLEQIASKDPST 290  
DB 175 VITREPKGYEAKKASVN-----PDIKVDYQAGSFDAKAGKTIIAAAMYAGADV 224  
OY 291 LLAVAGPL-TEIFSDIIA-----NONDR-YLIGVDTQSLV--YT---KTKNKEFTSL 337  
DB 225 VYQCAAGVGTVGFSEAKALNSTKNEADKRWYVIGVDDQGYLGGKYSKDGKDSNFVLVSTI 284  
OY 338 KNLGYSVFSVSLDLYTKKSNRNLGAFEPGKKSATVYLGINDRPDIADTSLSEGD 395  
DB 285 KEVG-----NVVKDIADKTKDGKFPGGTIVYDLKNGVNLGL-----DSANSEIK----- 330  
OY 396 LATEAISEAKKEFEKTKTIPAE 418  
DB 331 ---DAVAKAKADILDGKITVPSK 350

Search completed: July 18, 2002, 15:31:16  
Job time: 123 sec



**THIS PAGE BLANK (USPTO)**